

RL DNA RES; 5:203-216(1998).
DR EMBL; AB01390; BAB08575.1; -;
SEQUENCE 226 AA; 25019 MW; D8333100F9185B978 CRC64;

Query Match 100.0%; Score 29; DB 10;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Length 226;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
Db 196 SSGPSL 201

RESULT 9

ID	QBVK5	PRELIMINARY;	PRT;	257 AA.
AC	QBVK5;			
DT	01-MAR-2002 (TREMBLrel. 20; Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 22; Last annotation update)			
DB	Hypothetical protein MT0781.			
OS	Mycobacterium tuberculosis			
OC	Bacteria; Actinobacteria; Actinomycetales; Actinomycetales;			
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
OX	NCBI_TaxID=1773;			
RN	RN			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CDC 1551 / Oshkosh;			
RA	Fleischmann R.D., Allard D., Elsen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Giphian N.L., Haft D., Hickey B., Kolpogut J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khoruri H., Gill J., Mikula A., Bishai W.			
RA	"Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains"; Submitted (APR-2001) to the EMBL/GenBank/DDJB databases.			
RT	Submitted (APR-2001) to the EMBL/GenBank/DDJB databases.			
RL	EMBL; AB006999; AAK45022.1; -.			
DR	DR0781; -.			
KW	Hypothetical protein.			
SEQUENCE	257 AA; 25997 MW; 36D9A9AC973C5D54 CRC64;			
Query Match 100.0%; Score 29; DB 16; Best Local Similarity 100.0%; Pred. No. 1.5e-02; Length 257; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy 1 SSGPSL 6 Db 79 SSGPSL 84				

RESULT 10

ID	QBIXV0	PRELIMINARY;	PRT;	277 AA.
AC	QBIXV0;			
DT	01-MAR-2003 (TREMBLrel. 23; Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23; Last annotation update)			
DE	Similar to hairy homolog (Drosophila).			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Skin;			
RA	Streusberg R.; Submitted (NOV-2002) to the EMBL/GenBank/DDJB databases.			
RL	EMBL; BC039152; AAB39152.1; -.			
DR	GO; GO:0003677; F-DNA binding; IEA.			
GO	GO:0006335; P-regulation of transcription, DNA-dependent; IEA.			
DR	InterPro; IPR01052; HUH-basic.			
DR	InterPro; IPR003650; Orange.			

RESULT 11

ID	Q9L913	PRELIMINARY;	PRT;	284 AA.
AC	Q9L913;			
DT	01-OCT-2000 (TREMBLrel. 15; Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15; Last annotation update)			
DT	01-JUN-2003 (TREMBLrel. 24; Last annotation update)			
DE	COPB.			
GN	COPB.			
OS	Aeromonas veronii bv. sobria.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Aeromonadales;			
OC	Aeromonadaceae; Aeromonas.			
OX	NCBI_TaxID=11517;			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=3767;			
RA	Franci K.T., Chang B.J., Mee B.J., Collignon P.J., Keeese P.K.; "An adhesion defective mutant of Aeromonas veronii biovar sobria associated with copper tolerance"; Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.			
RT	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.			
RL	EMBL; AF187849; AAF12270.1; -.			
DR	IntePro; IPR007939; CopB.			
SQ	SEQUENCE 284 AA; 32401 MW; 68E414BE83DBF148 CRC64;			

RESULT 12

ID	Q8YJL9	PRELIMINARY;	PRT;	348 AA.
AC	Q8YJL9			
DT	01-MAR-2002 (TREMBLrel. 20; Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 20; Last annotation update)			
DE	Hypothetical membrane spanning protein BMF1063.			
GN	BMF1063.			
OS	Brucellimelitensis.			
OC	Bacteriaceae; Brucellaceae; Brucella.			
OX	NCBI_TaxID=29459;			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=16M / ATCC 23456 / Biotype 1;			
RX	DeAngelis V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T., Ivanova N., Anderson J., Bhattacharya A., Lykidis P., Reznik G., Jablonski L., Larsen N., Di Souza M., Bernal A., Mazur M., Gotsman E., Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J., Haselkorn R., Kyriakis N., Overbeek R.; "The Genome sequence of the facultative intracellular pathogen			

RT Brucella melitensis."; U.S.A. 99:443-448 (2002).
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
 DR EMBL: AE009450; ALU51245.1; -.
 DR FIR; AB3250; AB3260.
 DR InterPro: IPR008964; Invasin_intimin.
 KW Hypothetical protein; Complete Protein;
 SEQUENCE 348 AA; 2880442DA7B1774A CRC64;
 SQ Query Match 100.0%; Score 29; DB 16; Length 348;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSGPSSL 6
 Db 140 SSGPSSL 145

RESULT 13
 OTU35 PRELIMINARY; PRT; 358 AA.
 ID OTU35 PRELIMINARY;
 AC DT 01-OCT-2003 (TRMBLrel. 25, Created)
 DT 01-OCT-2003 (TRMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRMBLrel. 25, Last annotation update)
 DE Aminotransferases class-I (EC 2.6.1.9).
 DE Aminotransferase; Synecococcus sp. (strain WH8102).
 GN Synecococcus sp. (strain WH8102).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAUSBERG R.;
 RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RX MEDLINE=228569; PubMed=22917641;
 RA Palenik B.; Braamshaus B.; Larimer F.W.; Land M.; Hauser L.; Chain P.;
 RA Lamerdin J.; Regala W.; Allen E.E.; McCarren J.; Paulsen I.;
 RA Dufresne A.; Partensky F.; Webb E.A.; Waterbury J.;
 RT "The Genome of a motile marine Synechococcus";
 RL Nature 424:1037-1042(2003).
 DR BX56995; CAE05849.1; -.
 DR Aminotransferase; Transferase; Complete proteome.
 SQ SEQUENCE 358 AA; 40325 MW; OC35C403E54DC7F CRC64;

Query Match 100.0%; Score 29; DB 16; Length 358;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSGPSSL 6
 Db 92 SSGPSSL 97

Search completed: March 10, 2004, 09:25:29
 Job time : 8.92996 secs

RESULT 14
 Q7XK21 PRELIMINARY; PRT; 392 AA.
 ID Q7XK21 PRELIMINARY;
 AC DT 01-OCT-2003 (TRMBLrel. 25, Created)
 DT 01-OCT-2003 (TRMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRMBLrel. 25, Last annotation update)
 DE OSINBA0044K18.24 protein.
 GN OSINBA0044K18.24.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Enhaertcoideae; Oryzeae; Oryza.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fu G.; Wang S.Y.; Ren S.X.; Lv G.; Lin W.; Gu W.Q.; Zhu G.F.; Tu Y.F.;
 RA Jia J.; Yin H.F.; Zhang Y.; Cai Z.; Chen J.; Kang H.; Chen X.Y.;
 RA Shao Y.; Sun Y.; Hu Q.P.; Zhang X.L.; Zhang W.; Wang J.; Ding C.W.;
 RA Sheng H.H.; Gu J.L.; Chen S.T.; Ni L.; Zhu F.H.; Han B.; Feng Q.;
 RA Huang Y.C.; Li Y.; Zhu J.J.; Zhao Q.; Hu X.; Liu Y.L.; Mu J.; Yu Z.;
 RA Chen L.; Pan D.L.; Weng Q.J.; Zhang L.; Lu Y.Q.; Yu S.L.; Liu X.H.;
 RA Lu T.T.; Zhang Y.J.; Lu Y.; Li C.; Li T.; Zhang Y.; Hu H.; Jia P.X.,

GenCore version 5.1.6									
Copyright (c) 1993 - 2004 Compugen Ltd.									
Protein search, using SW model									
Run on: March 10, 2004, 08:58:53 ; Search time 1.09728 Seconds (without alignments) 284.724 Million cell updates/sec									
Title: US-09-848-834A-6									
perfect score: 29									
Sequence: 1 SSGPSL 6									
Scoring table: BLOSUM62									
Gapopen 10.0 , Gapext 0.5									
Searched: 141681 seqs, 52070155 residues									
Total number of hits satisfying chosen parameters: 141681									
Minimum DB seq length: 0									
Maximum DB seq length: 2000000000									
Post-processing: Minimum Match 0%									
Maximum Match 100%									
Listing First 45 summaries									
Database : SwissProt 42.2*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
%									
result Query Length DB ID Description									
No.	Score	Match	Length	DB	ID				
1	29	100.0	225	1	VV SV41	P36315 simian virus	Q14469 homo sapien	Q86606 simian virus	Q14469 homo sapien
2	29	100.0	280	1	HES1_HUMAN	Q86606 simian virus	Q14469 homo sapien	Q14469 homo sapien	Q14469 homo sapien
3	29	100.0	395	1	RRPP_SV41	Q14469 homo sapien	Q14469 homo sapien	Q14469 homo sapien	Q14469 homo sapien
4	29	100.0	410	1	HXA3_HETTER	Q14469 homo sapien	Q14469 homo sapien	Q14469 homo sapien	Q14469 homo sapien
5	29	100.0	416	1	NAP5_HUMAN	Q14469 homo sapien	Q14469 homo sapien	Q14469 homo sapien	Q14469 homo sapien
6	29	100.0	989	1	FRO_NEURR	P19970 neurospora	P09070 sordaria fi	P09070 sordaria fi	P09070 sordaria fi
7	29	100.0	997	1	PRO_SORF1	P09070 sordaria fi	P09070 sordaria fi	P09070 sordaria fi	P09070 sordaria fi
8	29	100.0	1149	1	DPO1_ADDCC	Q65946 canine aden	Q65946 canine aden	Q65946 canine aden	Q65946 canine aden
9	29	100.0	1150	1	DPOL_ADDCT	Q96kg7 homo sapien	Q96kg7 homo sapien	Q96kg7 homo sapien	Q96kg7 homo sapien
10	29	100.0	1210	1	BAT8_HUMAN	Q92148 mus musculu	Q92148 mus musculu	Q92148 mus musculu	Q92148 mus musculu
11	29	100.0	1263	1	BAT9_MOUSE	Q46974 saccharomyces cerevisiae	Q46974 saccharomyces cerevisiae	Q46974 saccharomyces cerevisiae	Q46974 saccharomyces cerevisiae
12	29	100.0	1380	1	ZNS1_YEAST	Q8sx83 drosophila melanogaster	Q9pt84 gallus gallus	Q9pt84 gallus gallus	Q9pt84 gallus gallus
13	29	100.0	5560	1	SPN_DRONE	Q9pt84 gallus gallus	Q9pt84 gallus gallus	Q9pt84 gallus gallus	Q9pt84 gallus gallus
14	27	93.1	526	1	KCH2_CHICK	Q95151 mycto mouse	Q95151 mycto mouse	Q95151 mycto mouse	Q95151 mycto mouse
15	27	93.1	102	1	MYCD_MOUSE	Q8r517 rattus norvegicus	Q8r517 rattus norvegicus	Q8r517 rattus norvegicus	Q8r517 rattus norvegicus
16	27	93.1	938	1	MYCD_RAT	Q9er47 mus musculus	Q9er47 mus musculus	Q9er47 mus musculus	Q9er47 mus musculus
17	27	93.1	1195	1	KCH7_MOUSE	Q54852 rattus norvegicus	Q54852 rattus norvegicus	Q54852 rattus norvegicus	Q54852 rattus norvegicus
18	27	93.1	1195	1	KCH7_RAT	Q9ns40 homo sapien	Q9ns40 homo sapien	Q9ns40 homo sapien	Q9ns40 homo sapien
19	27	93.1	1196	1	KCH7_HUMAN	Q9uhv7 homo sapien	Q9uhv7 homo sapien	Q9uhv7 homo sapien	Q9uhv7 homo sapien
20	27	93.1	2174	1	T240_HUMAN	P00076 euglena gracilis	P06892 xenopus laevis	P06892 xenopus laevis	P06892 xenopus laevis
21	26	89.7	102	1	CIC_EUGGER	Q93591 caenorhabditis elegans	Q93591 caenorhabditis elegans	Q93591 caenorhabditis elegans	Q93591 caenorhabditis elegans
22	26	89.7	209	1	HIA_XENIA	Q9xsk0 bovis taurinus	Q9xsk0 bovis taurinus	Q9xsk0 bovis taurinus	Q9xsk0 bovis taurinus
23	26	89.7	297	1	YUBO_CABEL	Q98737 kielbisia obovata	Q98737 kielbisia obovata	Q98737 kielbisia obovata	Q98737 kielbisia obovata
24	26	89.7	299	1	CRX_BOVIN	P057451 mus musculus	P057451 mus musculus	P057451 mus musculus	P057451 mus musculus
25	26	89.7	299	1	CRX_HUMAN	P06712 merluccius productus	P06712 merluccius productus	P06712 merluccius productus	P06712 merluccius productus
26	26	89.7	299	1	CRX_MOUSE	P07544 bacillus caldotenuis	P07544 bacillus caldotenuis	P07544 bacillus caldotenuis	P07544 bacillus caldotenuis
27	26	89.7	366	1	PRB_MEETH	Q92040 paramyxo virus	Q92040 paramyxo virus	Q92040 paramyxo virus	Q92040 paramyxo virus
28	26	89.7	392	1	LEU3_BACCA	PF03210 paramyxo virus	PF03210 paramyxo virus	PF03210 paramyxo virus	PF03210 paramyxo virus
29	26	89.7	393	1	PSMR_SUTTO	Q96572 hordeum vulgare	Q96572 hordeum vulgare	Q96572 hordeum vulgare	Q96572 hordeum vulgare
30	26	89.7	393	1	PSMR_SUTSO	P21616 phascolus acutus	P21616 phascolus acutus	P21616 phascolus acutus	P21616 phascolus acutus
31	26	89.7	399	1	HXB3_CHICK	P31414 arabitidopsis thaliana	P31414 arabitidopsis thaliana	P31414 arabitidopsis thaliana	P31414 arabitidopsis thaliana
32	26	89.7	457	1	NIFE_KLBPN	P97680 rattus norvegicus	P97680 rattus norvegicus	P97680 rattus norvegicus	P97680 rattus norvegicus
33	26	89.7	496	1	NIFE_RHINO	P55673 rhizobium sp.	P55673 rhizobium sp.	P55673 rhizobium sp.	P55673 rhizobium sp.

HES1 HUMAN STANDARD; PRT; 280 AA.

ID HES1_HUMAN; STANDARD; PRT; 280 AA.

AC Q14479; (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Transcription factor HES-1 (Hairy and enhancer of split 1) (Hairy-like)

DE HES1 OR HRY (OR HHL).

GN Homo sapiens (Human).

OS Homo sapiens (Human).

OC Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CX NCBI_TAXID=9606;

RN [1]

SEQUENCE FROM N.A.

RA MEDLINE=94292187; PubMed=8020957;

RA Feder J.N., Li L., Jan L.Y., Jan Y.-N.;

RT "Genomic cloning and chromosomal localization of HRY, the human homolog to the Drosophila segmentation gene, hairy.";

RL Genomics 20:56-61(1994).

[2]

SEQUENCE FROM N.A.

RA TISSUE=Kidney;

RA Yao J., Yeung S., Sun H., Chen N.;

RT "Functional analysis of human HRY in Drosophila.";

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ/NCBI databases.

[3]

RN

SEQUENCE FROM N.A.

RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T., Nakamura Y., Obayashi T., Shiba T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;

RA Nakamura Y., Isogai T., Sugano S.;

RT "NEO human cDNA sequencing Project.";

RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ/NCBI databases.

[4]

RN

INTERACTION WITH SIRT1.

RA MEDLINE=22422390; PubMed=12555671;

RT "Human Sir2-related protein SIRT1 associates with the bHLH repressors HES1 and HEY2 and is involved in HES1- and HEY2-mediated transcriptional repression.";

RT Biochem Biophys Res Commun. 301:250-257 (2003).

CC -!- FUNCTION: transcriptional repressor of genes that require a bHLH protein for their transcription. May act as a negative regulator of myogenesis by inhibiting the functions of MYOD1 and ASH1 (BY similarity).

CC -!- SUBUNIT: transcription repression requires formation of a complex with a co-repressor protein (Groucho) such as TLE2 (BY similarity). Interacts with SIRT1.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- DOMAIN: Has a partial type of basic domain (presence of a helix-interrupting proline) that binds to the N-box (CACNAG), rather than the canonical E-box (CANNTG).

CC -!- DOMAIN: The carboxy-terminal WRW motif is a transcriptional repression domain necessary for the interaction with GROUCHO, a transcriptional co-repressor recruited to specific target DNA by Hairy-related proteins' basic helix-loop-helix (bHLH) domain.

CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

CC -!- SIMILARITY: Contains 1 orange domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR MIM; 139605; -; F:DNA binding; TAS.

DR GO; GO:0003677; F:DNA binding; TAS.

DR GO; GO:0007399; F:neurogenesis; TAS.

DR InterPro; IPR01692; HLH basic.

DR InterPro; IPR03550; Orange.

DR Pfam; PF00010; HLH; 1.

DR SMART; SM00353; HLH; 1.

DR PROSITE; PS5588; HLH; 1.

DR Transcription regulation; DNA-binding; Nuclear protein; Repressor; BASIC DOMAIN.

KW DNA BIND 35 47

FT DNA BIND 35 47

FT DOMAIN 48 92

FT DOMAIN 156 246

FT DOMAIN 249 273

FT DOMAIN 275 278

FT DOMAIN 280 AA; 28541 MW; P9342488FC749B3C C9064;

SO Query Match 100.0%; Score 29; DB 1; Length 280;

Best Local Similarity 100.0%; Pred. No. 37;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPGL 6

Db 264 SSGPGL 269

RESULT 3

RRPP SV41 STANDARD; PRT; 395 AA.

ID RRPP_SV41

AC Q86506; (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

RNA Polymerase alpha subunit (EC 2.7.7.48) (P protein).

DN P/V

GN Simian virus 41 (SV41).

OS ssRNA negative-strand viruses; Mononegavirales;

OC Paromyxovirinae; Paramyxoviridae; Rubulavirus.

OC

OX NCBI_TaxID111228;

RN

SEQUENCE FROM N.A. AND RNA EDITING.

RC STRAN=Toshiba/Charnock;

RC MEDLINE=93260108; PubMed=8492098;

RX RA Kawano M., Tarudome M., Oki N., Nishio M., Komada H., Matsumura H., Kusagawa S., Ohta H., Ito Y.;

RA "Sequence determination of the P gene of simian virus 41: presence of irregular deletions near the RNA-editing sites of paramyxoviruses.";

RT RT irregular deletions near the RNA-editing sites of paramyxoviruses.";

RT RL J. Gen. Virol. 74:911-916 (1993).

CC -!- FUNCTION: Essential component of the RNA polymerase and the nascent chain assembly complex. Also required during RNA synthesis.

CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA) (N).

CC -!- RNA EDITING: Modified positions=16; Note=Partially edited. RNA editing at this position consists of an insertion of two guanine nucleotides. The sequence displayed here is the P protein, derived from the edited RNA. The unedited RNA version gives rise to the V protein (AC P36315).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC --- S60813; AAB26640; 2;

DR EMBL; S60813; AAB26640; 2;

DR PIR; JG2041; JG2041;

DR InterPro; IPR004897; Paramyx_P_V.

DR Pfam; PF03210; Paramyx_P_V_1.

KW Transferase; RNA-directed RNA Polymerase; Nucleocapsid;

DR Genew; HGNC:5192; HES1.

KW Phosphorylation; RNA editing;
SEQUENCE 395 AA; 41920 MW; 6897477FBF483D30 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 395;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
DB 144 SSGPSL 149

RESULT 4
HXA3_HETFR STANDARD PRT; 410 AA.

ID HXA3_HETFR
AC Q9IA21;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Homeobox protein Hox-A3.
GN HOXA3.
OS Heterodontus francisci (Bony shark).
OC Blasmodibranchii; Chordata; Craniata; Vertebrata; Chondrichtyes;
OC Heterodontidae; Heterodontidae; Heterodontiformes;
OC NCBI_TaxID=7792;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20144096; PubMed=10677514;
RA Kim C.B., Amemiya C., Bailey W., Kawasaki K., Mezey J., Miller W.,
RA Minoshima S., Shimizu N., Wagner G., Ruddle F.;
PT "Hox cluster genetics in the horn shark, *Heterodontus francisci*."
RL Proc. Natl. Acad. Sci. U.S.A. 97:1655-1660 (2000).
CC !-- PRODUCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis (By
CC SIMILARITY).
CC !-- SUBCELLULAR LOCATION: Nuclear (By Similarity).
CC !-- SIMILARITY: Belongs to the Antp homeobox family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; AP224262; AAF44641.1; -.
CC HSSP; P02833; ISSN.
DR TRANSFIC; T04474; -.
DR InterPro; IPR001527; Antennapedia.
DR InterPro; IPR001556; Homeobox.
DR PRINTS; PR00025; ANTEENNAPEDIA.
DR PRINTS; PR00024; HOMEOBOX.
DR PRODOM; PD00010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00032; ANTEENNAPEDIA; 1.
DR PROSITE; PS00027; HOMEOBOX; 1.
DR PROSITE; PS50071; HOMEOBOX; 2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT SITE 133 138 ANTP-TYPE HEXAPEPTIDE.
FT DOMAIN 168 227 HOMEOBOX.
FT DOMAIN 85 90 POLY-GLN.
FT DOMAIN 91 100 POLY-PRO.
FT DOMAIN 149 153 POLY-SER.
SQ SEQUENCE 410 AA; 44548 MW; 285ABC06BA1C9FD9 CRC64;

QY 1 SSGPSL 6
DB 346 SSGPSL 351

RESULT 5
NAP5_HUMAN STANDARD PRT; 416 AA.

ID NAP5_HUMAN
AC O14515;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Nck-associated Protein 5 (NAP-5) (Fragment).

GN NAP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE:Brain;
RX MEDLINE=98008866; PubMed=9344857;
RA Matuoka K., Miki H., Takahashi K., Takenawa T.;
RT "A novel ligand for an SH3 domain of the adapter protein Nck bears an
RT SH2 domain and nuclear signaling motifs";
RL Biochem. Biophys. Res. Commun. 239:488-492 (1997).
CC !-- SUBUNIT: Interacts with the SH3-containing region of the adapter
CC protein Nck.
CC !-- TISSUE SPECIFICITY: Expressed in fetal and adult brain, leukocytes
CC and fetal fibroblasts.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC DR EMBL; AB005217; BAA22433.1; -.
CC FT NON_TER 416 416 AE4B7F3A19522612 CRC64;
CC SQ SEQUENCE 416 AA; 44829 MW; AE4B7F3A19522612 CRC64;
Query Match 100.0%; Score 29; DB 1; Length 416;
Best local similarity 100.0%; Pred. No. 55;
Matches 6; conservative 0; mismatches 0; indels 0; gaps 0;

QY 1 SSGPSL 6
DB 192 SSGPSL 197

RESULT 6
FRO_NEUCR STANDARD PRT; 989 AA.

ID FRO_NEUCR
AC P19370; Q01276;
DT 01-FEB-1991 (Rel. 17, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Frequency clock protein.

GN FRO
OS Neurospora crassa.
OC Sordariomycetidae; Sordariales; Sordariaceae; Sordariomyces;
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomyces;
NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
STRAIN=74-OR3-1A / FGSC 987;
RC MEDLINE=94325580; PubMed=8032643;
RX Aronson B.D.; Johnson K.A.; Dunlap J.C.;
RA Aronson B.D.; Johnson K.A.; Dunlap J.C.;
RT "Circadian clock locus frequency; protein encoded by a single open

Query Match 100.0%; Score 29; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

reading frame defines Period length and temperature compensation.";

Proc. Natl. Acad. Sci. U.S.A. 91:7683-7687(1994).

[2] SEQUENCE OF 202-989 FROM N.A.
MEDLINE=89241721; PubMed=255233;
McClung C.R.; Fox B.A.; Dunlap J.C.;
"The Neurospora clock gene frequency shares a sequence element with
the Drosophila clock gene period.";
Nature 339:558-562(1998).

[3] PHOSPHORYLATION BY CK2.
MEDLINE=21955574; PubMed=11959847;
Yang Y.; Cheng P.; Liu Y.;
"Regulation of the Neurospora circadian clock by casein kinase II.";
Genes Dev. 16:994-1006(2002).

-!- FUNCTION: Circadian clock component involved in the generation of biological rhythms, in particular in rhythm stability, period length, and temperature compensation. Oscillates in abundance with a daily peak early in the morning. Behaves as a negative element in circadian transcriptional loop. May bind to wc2 protein. The complex frg-wc2 may turn off the expression of frg.

-!- ALTERNATIVE PRODUCTS:
Event=Alternative initiation;
Comment=2 isoforms. Long (shown here) and short, are produced at alternative initiation sites. The isoform long maintains rhythms at high temperature (30 degrees Celsius), while the isoform short maintains rhythms at lower temperature (18 degrees Celsius);

-!- INDUCTION: By light; perhaps through white collar-1 (wc1) and white collar-2 (wc2). Also activated directly by wc1 and wc2.

-!- PTM: Progressive phosphorylation during the late circadian day early night. Phosphorylation is also involved in regulating frg degradation. Phosphorylation by CKII may have at least three functions; it decreases the stability of frg, reduces the probe complex formation between frg and the white collar proteins, and is important for the closing of the Neurospora circadian negative feedback loop.

RESULT 7

FRQ_SORPI	STANDARD;	FRT;	997 AA.
ID_FRO_SORPI			
AC_Q01033;			
DT 15-JUL-1999	(Rel. 38, Created)		
DT 15-JUL-1999	(Rel. 38, Last sequence update)		
DT 28-FEB-2003	(Rel. 41, Last annotation update)		
DE Frequency clock protein.			
GN FRQ			
OS <i>Sordaria fimicola</i> . Fungi; Ascomycota; Pezizomycotinae; Eukaryota; Fungi; Ascomycota; Pezizomycotinae; Sordariaceae			
OC			
OC			
NCBI_TaxID=27338;			
OX [1]			
RN	SEQUENCE FROM N.A.		
RP			
RC STRAIN=FSGC 2918;			
RA MEDLINE=9425331; PubMed=8194516;			
RA Merrow M.W.; Dunlap J.C.;			
RT "Intergeneric complementation of a circadian phylogenetic conservation of structure and frequency."			
RT			
RL EMBO J. 13:2257-2266 (1994).			
CC -1- FUNCTION: Circadian clock component involved in biological rhythms, in particular in rhythmic length, and temperature compensation. Belongs to CC element in circadian transcriptional loop.			
CC -!- SIMILARITY: BELONGS TO THE FRQ FAMILY.			
CC			
CC This SWISS-PROT entry is copyright. It is provided by the Swiss Institute of Bioinformatics between the European Bioinformatics Institute. Their use by non-profit institutions as long as modified and this statement is not removed. See below for details.			
CC or send an email to license@ibb-sib.ch.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial requirements requires a licence agreement (See <http://www.ebi.ac.uk/announce/>) or send an email to license@ebi.ac.uk.

Best Local Similarity 100.0% ; Pred. No. 1.e+02 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;							
Matches 6 ; Conservative 0 ; SSGPSL 282							
Qy 1 SSGPSL 6 Db 277 SSGPSL 282							
RESULT 8							
DPO1_ADECC	ID	DPO1_ADECC	STANDARD;	PRT;	1149 AA.		
065346;	AC						
01-NOV-1997	DT	35; Created					
01-NOV-1997	DT	(Rel. 35, Last sequence update)					
28-FEB-2003	DT	(Rel. 41, Last annotation update)					
DNA polymerase	DE	(EC 2.7.7.7).					
GN	GN	Canine adenovirus type 1 (strain CL1).					
OC	OC	Viruses; dsDNA viruses; no RNA stage; Adenoviridae; Mastadenoviruses.					
OX	OX						
NCBI_TaxID=69150;	RN	[1]					
SEQUENCE FROM N.A.	RP						
CAMPBELL J.B.; ZHAO Y.;	RA						
Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.	RL						
-1- CATALYTIC ACTIVITY: N deoxyribonucleoside triphosphate = N diphosphate	CC						
+ [DNA] (N).	CC						
-1- MISCELLANEOUS: This DNA Polymerase requires a protein as a primer.	CC						

```

Query Match      100.0%; Score 29; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches       6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy          1 SSCPSSL 6
Db          271 SSCPSSL 276

```

OC Viruses; dsDNA viruses; no RNA viruses; Archaea; Bacteria; eukaryotes.

NCBI_TAXID=69150;

[1] RN

SEQUENCE FROM N.A.

RP Campbell J.B., Zhao Y.;

RA Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.

RL -1 - CATALYTIC ACTIVITY: N deoxyribonucleoside triphosphate = N diphosphate

CC + [DNA] (N)

CC -1 - MISCELLANEOUS: This DNA polymerase requires a protein as a primer.

CC

-!- SIMILARITY: Belongs to the DNA polymerase type-B family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U55001; AAB05134; 1; -;
 DR InterPro; IPR006172; DNA_Pol_B_1.
 DR InterPro; IPR004868; DNA_Pol_B_2.
 DR InterPro; IPR004137; Prismane_Pfam; PF03175; DNA_Pol_B_1.
 DR PRINTS; PRO0106; DNAPOLB.
 DR SMART; SM00486; POLB1.
 DR PROSITE; PS00116; DNA POLYMERASE_B_1.
 DR Transferase; DNA-directed DNA Polymerase; DNA replication;
 KW DNA-binding.
 SQ SEQUENCE 1149 AA; 130393 MW; A0B686PAAFFPA236E CRC64;
 Query Match 100.0%; Score 29; DB 1; Length 1149;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SSGPSSL 6
 Db 25 SSGPSSL 30

RESULT 10

BAT8_HUMAN STANDARD PRT; 1210 AA.
 ID BAT8_HUMAN STANDARD PRT; 1210 AA.
 AC Q96K07; Q14349; Q96M95; Q96QD0; Q9uQL8; Q9Y331;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Histone-Lysine N-methyltransferase H3 Lysine-9 Specific 3
 DE (EC 2.1.1.43) (Histone H3-K9 methyltransferase 3) (H3-K9-HMTase 3)
 DE (HLA-B associated transcript 8) (G9a) (NG36).
 GN BAT8 OR G9A OR NC36.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBITaxonID=9606;
 OX
 RN [1] SEQUENCE FROM N.A.
 RA Rown L., Qin S., Madan A., Dickhoff R., Dors M., Madan A., Hicks P., Rowen L., Qin S., Madan A., Dickhoff R., Dors M., Madan A., Hicks P.,
 RA Loerz C., Ratcliffe A., Abbas N., Shaffer T., Hood L., Hood L.,
 RT "Sequence of the human major histocompatibility complex class III region";
 RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RL [2] SEQUENCE FROM N.A.
 RA Hirakawa M., Yamaguchi H., Imai K., Shimada J., Shiina S., Tamiya G., Oka A., Inoko H.;
 RA "Homo sapiens 2"; 817bp genomic DNA of 6p21.3 HLA class I region."
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RL [3] SEQUENCE FROM N.A. (ISOFORM 1), ALTERNATIVE SPlicing (ISOFORM 2), AND TISSUE SPECIFICITY.
 RN [4] SEQUENCE FROM N.A.
 RA Brown S.B., Campbell R.D., Sanderson C.M.;
 RA "Novel NG36/G9a gene products encoded within the human and mouse MHC class III regions";
 RT PubMed=21564388; PubMed=11707778;
 RL [5] SEQUENCE FROM N.A. (ISOFORM 1).
 RN [4] SEQUENCE FROM N.A. (ISOFORM 3).
 RP TISSUE=Salivary Gland;
 RC Oshima A., Takahashi-Puji A., Tanase T., Imose N., Takeuchi K., Arita M., Musashino K., Yunki H., Sugiyama T., Irie R., Otsubo T., Matsumoto A., Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H., Matsuo K., Nakamura Y., Seike M., Kikuchi H., Kaneko B., Wagatsuma M., Murakawa K., Kaneko K., Kawakami B., Suzuki Y., Sugano S., Nagahari K., Matsuno Y., Nagai K., Tsogai T.;
 RT "NEO human cDNA sequencing Project";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [5] SEQUENCE OF 5-1210 FROM N.A. (ISOFORM 1).
 RP TISSUE=Muscle, and Uterus;
 RC MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K., Altenschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Tohidiyuki S., Carninci P., Prange C., Rana S.S., Logueillo N.M.A., Peters G.J., Abramson R.D., Mullahay S.J., Borsig S.A., McEwan P.J., McKernan K.J., Malek J.A., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Richards S., Worley K.C., Villaillon D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahney J., Helton B., Ketteman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.; Rodriguez A.C.; Grinberg J.; Schmitz J.; Myers R.M.; Krzywinski M.I.; Skalska U.; Smailus D.E.; Schnera A.; Schein J.B.; Jones S.J.M.; Marra M.A.; RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:1689-16903 (2002).
RN [6]
RP SEQUENCE OP 195-1210 FROM N.A., AND CHARACTERIZATION.
RC TISSUE-Histiocytic lymphoma;
RX MEDLINE#93207535; PubMed#1457211;
RA Milner C.M.; Campbell R.D.;
RT "The G9a gene in the human major histocompatibility complex encodes a novel protein containing ankyrin-like repeats.";
RL Biochem. J. 290:811-818 (1993).
RN [7]
RP IDENTIFICATION IN COMPLEX WITH L3MBTL2 AND YAF2.
RX MEDLINE#21326082; PubMed#11316813;
RA Tachibana M.; Sugimoto K.; Fukushima T.; Shinkai Y.; Set domain-containing protein, G9a, is a novel lysine-prefering mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3.";
RL J. Biol. Chem. 276:25309-25317 (2001).
RN [8]
RP IDENTIFICATION IN COMPLEX WITH L3MBTL2 AND YAF2.
RX MEDLINE#21995559; PubMed#12004135;
RA Ochiai H.; Ishiguro K.-I.; Saubatz S.; Livingston D.M.; Nakatani Y.; RT "A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells.";
RL Science 296:112-1136 (2002).
CC -1- FUNCTION: Histone methyltransferase. Preferentially methylates Lys-9 of histone H3 and Lys-27 of histone H3 (in vitro); H3 Lys-9 methylation represents a specific tag for epigenetic transcriptional repression by recruiting HPI proteins to methylated histones. Probably targeted to histone H3 by different DNA-binding proteins like E2F6, MAX, MAX and/or DP1. Also methylates histone H1 (by similarity).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine = S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
CC -1- SUBUNIT: Part of the E2F6 com-1 complex in G0 phase composed of E2F6, MAX, YAF2, L3MBTL2, RNF1, RNF2, MBLR, L3MBTL2 and YAF2.
CC -1- SUBCELLULAR LOCATION: Nuclear; associates with euchromatic regions. Does not associate with heterochromatin.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms;
CC Comment=Additional isoforms seem to exist;
CC Name=-;
CC IsoId=Q96KQ7-1; Sequence=Displayed;
CC IsoId=Q96KQ7-2; Sequence=VSP_002211;
CC Name=-; Synonyms=NG36G9a; SPI=2213;
CC IsoId=Q96KQ7-3; Sequence=VSP_002212; VSP_002213;
CC -1- TISSUE SPECIFICITY:Expressed In all tissues examined, with high levels in fetal liver, thymus, lymph node, spleen and peripheral blood leukocytes and lower level in bone marrow.
CC -1- SIMILARITY: Belongs to the histone-lysine methyltransferase family. Suvar3-9 subfamily.
CC -1- SIMILARITY: Contains 7 ANK repeats.
CC -1- SIMILARITY: Contains 1 pre-SET domain.
CC -1- SIMILARITY: Contains 1 SET domain.
CC -1- SIMILARITY: Contains 1 post-SET domain.
CC -1- CAUTION: NC36 and G9a were originally thought to derive from two separate genes. Ref.2 shows that all G9a transcripts also contain the in frame coding sequence of NG36.
CC -1- CAUTION: Ref.1 and Ref.2 sequences differ from that shown due to erroneous gene model prediction.
CC -1- CAUTION: It is not known whether Met-1 or Met-21 is the initiator methionine.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC DR EMBL; AF134726; ADD2811.1; ALT SEQ;
CC DR EMBL; AF134726; ADD2812.1; ALT SEQ;
CC DR EMBL; AP000502; BAB6395.1; ALT SEQ;
CC DR EMBL; AP000502; BAB6394.1; ALT SEQ;
CC DR EMBL; AJ315532; CAC8666.1; ALT SEQ;
CC DR EMBL; AK056336; BAB71314.1; ALT INIT;
CC DR EMBL; BC002586; AAH02686.1; ALT INIT;
CC DR EMBL; BC018718; AAH1818.1; ALT INIT;
CC DR EMBL; BC020970; AAH20970.1; ALT INIT;
CC DR EMBL; X69838; CAR49401.1; ALT INIT;
CC DR Genew; HGNC:14123; BAT8.
CC MIM: 604599; -
CC DR InterPro; IPR002110; ANK;
CC DR InterPro; IPR003616; PostSET;
CC DR InterPro; IPR007728; Pre-SET;
CC DR InterPro; IPR001214; SET;
CC DR InterPro; IPR003406; Zn2-binding.
CC DR Pfam; PF00023; anki; 6.
CC DR Pfam; PF05033; Pre-SET; 1.
CC DR Pfam; PF0855; SET; 1.
CC DR SMART; PRO1415; ANKYRIN.
CC SMART; SMC0248; ANK; 6.
CC SMART; SM00458; PRESET; 1.
CC DR PROSITE; PS50088; ANK_REPEAT; 5.
CC DR PROSITE; PS50297; ANK REP REGION; 1.
CC DR PROSITE; PS50868; POST_SET; FALSE_NEG.
CC DR PROSITE; PS50867; PRE_SET; 1.
CC DR PS50280; SET; 1.
CC KW Transferase; Methyltransferase; Chromatin regulator; Nuclear protein; KW Transferase; Methionyltransferase; Chromatin regulator; Nuclear protein; KW Transferase; Alternative splicing.
CC FT DOMAIN 1 163
CC FT DOMAIN 2 300 326
CC FT DOMAIN 3 649 678
CC FT REPEAT 684 713
CC FT REPEAT 717 746
CC FT REPEAT 750 780
CC FT REPEAT 784 813
CC FT REPEAT 817 846
CC FT REPEAT 850 879
CC FT DOMAIN 4 972 1035
CC FT DOMAIN 5 1037 1159
CC FT DOMAIN 6 1164 1180
CC FT VARSPLIC 373 406
CC FT VARSPLIC 195 202
CC FT VARSPLIC 203 1210
CC FT VARSPLIC 203 Missing (In isoform 3).
CC FT VARSPLIC 203 /FTDID=VSP_002212.
CC FT VARSPLIC 203 Missing (In isoform 3).
CC FT VARSPLIC 203 /FTDID=VSP_002213.
CC FT CONFLICT 5 5
CC FT CONFLICT 13 13
CC FT CONFLICT 55 55
CC FT CONFLICT 168 169
CC FT CONFLICT 178 178
CC FT CONFLICT 985 985
CC FT CONFLICT 994 994
CC SEQUENCE 1210 AA; 132383 MW; B8508923CA6911A CRC64;
CC Query Match 100.0%; Score 29; DB 1; Length 1210;
CC Best Local Similarity 100.0%; Pred. No. 1.7e+02; Gaps 0;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC Qy 1 SSGPSL 6
CC Db 605 SSGPSL 610

RESULT 11

BATS MOUSE STANDARD; PRT; 1263 AA.
ID BAT8_MOUSE_08K4R6; Q8Z149;
AC Q9Z148;
DT 28-FEB-2003 (Rel. 41, Create)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DT Histone-Lysine N-methyltransferase, H3 Lysine-9 specific 3 (H3-K9-HMTase 3)
DB (EC 2.1.1.43) (Histone H3-K9 methyltransferase 3) (H3-K9-HMTase 3)
DE (HLA-B associated transcript 8) (G9a) (NG36).
GN BATS OR G9A OR NG36.
OS Mus musculus (Mouse).
OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Buteleostomi;
OX NCBI_TAXID=10090; Name=G9a; Name=;

RN [1] SEQUENCE FROM N.A.
RC STRAIN=129;
RA Schaffer T., Madan A., Qin S., Dahl T., James R., Dickhoff R.,
RA Rowen L., Madan A., Qin S., Hall J., Dahl T., James R., Dickhoff R.,
RA Schaffer T., Ratcliffe A., Abbasi N., Loretz C., Lasry S., Hood L.;
RT "Sequence of the mouse MHC class III region";
RL Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN [2] SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND MUTAGENESIS OF
RP 1165 ASN-CYS-1168.
RX MEDLINE=22123403; PubMed=1210538;
RA Tachibana M., Sugimoto K., Nozaki M., Ueda J., Ohta T., Ohki M.,
RA Fukuda M., Takeo N., Nida H., Kato H., Shinkai Y.;
RT "G9a histone methyltransferase plays a dominant role in euchromatic histone H3 Lysine 9 methylation and is essential for early embryogenesis";
RT Genes Dev. 16:1779-1791 (2002).
RN [3] SEQUENCE OF 960-1263 FROM N.A.
RC TISSUE=Breast tumor;
RX MEDLINE=2338825; PubMed=12477932;
RA Straubhaar R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
PA Klausner R.D., Collins F.S., Wagner L., Sheppard C.M., Schuler G.D.,
RA Hopkins R.F., Jordaan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soriano M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rosa S.A., Loqueland N.A., Peters G., Abramson R.D., Mulahay S.J.,
RA Richards S., McElvan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kertesz M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shervchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinbaum J., Schmitz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherzer A., Schein J.E., Jones S.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4] ALTERNATIVE SPLICING (ISOFORM 2).
RX MEDLINE=21564388; PubMed=11707778;
RA Brown S.E., Campbell R.D., Sanderson C.M.;
RT "Novel NG36/G9a gene products encoded within the human and mouse MHC class III regions";
RN Mamm. Genome 12:916-924 (2001).

RN [5] ENZYMATIC ACTIVITY, SUBCELLULAR LOCATION, AND MUTAGENESIS OF ARG-1162.
RX MEDLINE=21326082; PubMed=11316813;
RA Tachibana M., Sugimoto K., Fukushima T., Shinkai Y.;
RT "Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selection to lysines 9 and 27 of histone H4.";
RT J. Biol. Chem. 276:25309-25317 (2001).
RT -I: FUNCTION: Histone methyltransferase. Preferentially methylates

Lys-9 of histone H3 and Lys-27 of histone H3 (in vitro). H3 Lys-9 methylation represses a specific tag for epigenetic transcriptional repression by recruiting HPI proteins to methylated histones. Also methylates histone H1. CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine = S-adenosyl-L-homocysteine + histone N-(6)-methyl-L-lysine. SUBUNIT: Part of the E2F6.com1 complex in G0 phase composed of E2F6, MGA, MAX, CBX3, BAT8, EUMASAE1, RING1, RNF2, MBLR, L3MBTL2 and YAF2 (By similarity). SUBCELLULAR LOCATION: Nuclear. almost excluded from nucleoli. Associates with euchromatic regions. Does not associate with heterochromatin.

-I: ALTERNATIVE PRODUCTS:
Event-Alternative Splicing; Named isoforms=2;
Event-Alternative Splicing; Named isoforms=G9a-L;
Name=G9a-S;
Name=G9a-L;
IsoID=Q9Z148-1; Sequence=VSP_002215; VSP_002216;
IsoID=Q9Z148-2; Sequence=VSP_002214; VSP_002216;
-I: MISCELLANEOUS: G9a deficient mice show a higher level of histone H3 with acetylated Lys-9 and/or methylated Lys-4, display severe developmental defects and die within E9.5-E12.5 stages. SIMILARITY: Belongs to the histone-lysine methyltransferase family. Suvar3-9 subfamily.

-I: SIMILARITY: Contains 7 ANK repeats.

-I: SIMILARITY: Contains 1 SET domain.

-I: SIMILARITY: Contains 1 SET domain.

-I: SIMILARITY: Contains 1 post-SET domain.

-I: SIMILARITY: Contains 1 post-SET domain.

-I: SIMILARITY: Contains 1 SET domain.

-I: SIMILARITY: Contains 1 SET domain.

-I: SIMILARITY: Contains 1 SET domain.

-I: CAUTION: Ref 1 sequences differ from that shown due to erroneous gene model prediction.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation [3]. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/CC or send an email to license@isb-sib.ch).

CC EMBL; AF109906; AAC81164.1; ALT_SEQ.
DR EMBL; AB109806; AAC81165.1; ALT_SEQ.
DR EMBL; AB077209; BAC05482.1; -.
DR EMBL; AB077210; BAC05483.1; -.
DR EMBL; AB077211; BAC05484.1; -.
DR MCP; MG1:2148922; Bat8.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR003616; Post-SET.
DR InterPro; IPR007728; Pre-SET.
DR InterPro; IPR001214; SET.
DR InterPro; IPR003606; Zn2+-binding.
DR SMRT; SM00468; PreSET; 1.
DR SMRT; SM00468; PreSET; 1.
DR PFAM; PF00043; ankyrin repeat; 6.
DR PFAM; PF05033; Pre-SET; 1.
DR PFAM; PF00056; SET; 1.
DR PRINTS; PRO1415; ANKTRIN.
DR SMRT; SM00248; ANK; 6.
DR SMRT; SM00468; PreSET; 1.
DR SMRT; SM00317; SET; 1.
DR PROSITE; PS51088; ANK REPEAT; 5.
DR PROSITE; PS50297; ANK REG; REGION; 1.
DR PROSITE; PS50069; POST_SET; FALSE_NEG.
DR PROSITE; PS51867; PRE_SET; 1.
DR PROSITE; PS50280; SET; 1.
KW Transferase; Methyltransferase; Chromatin regulator; Nuclear protein; ANK repeat; Alternative splicing.
FT DOMAIN 352 379 Poly-Glu.
FT REPEAT 702 731 ANK 1.
FT REPEAT 737 766 ANK 2.
FT REPEAT 770 799 ANK 3.
FT REPEAT 803 833 ANK 4.
FT REPEAT 837 866 ANK 5.
FT REPEAT 870 899 ANK 6.
FT REPEAT 903 932 ANK 7.

PT DOMAIN 1025 1088 PRE-SET.
 PT DOMAIN 1090 1212 SET-BET.
 PT DOMAIN 1217 1233 POST-BET.
 PT VARSPLIC 1 57 Missing (in isoform 2).
 /FTID=VSP_002214.
 AGLIGGPPVPCDPSQ -> MAAAGGAAAAAE (in
 isoform 2).
 /FTID=VSP_002215.
 Missing (In isoform 2).
 /FTID=VSP_002216.
 R->H: STRONGLY REDUCES HISTONE
 METHYLTRANSFERASE ACTIVITY.
 PT MUTAGEN 1162 1162 MISSING: ABOLISHES HISTONE
 METHYLTRANSFERASE ACTIVITY AND SUBSEQUENT
 PT MUTAGEN 1165 1168 PREDICTION.
 PT SEQUENCE 1263 AA; 138038 MW; 74DBFF9A3679589 CRC64;
 Query Match 100.0%; Score 29; DB 1; Length 1263;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;
 QY 1 SSGPSSL 6
 Db 658 SSGPSSL 663

RESULT 12
 ZMS1_YEAST STANDARD; PRT; 1380 AA.
 ID ZMS1_YEAST STANDARD; PRT; 1380 AA.
 AC P4674;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Zinc finger protein ZMS1.
 DE ZMS1 OR YJR127C OR J2012.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OC NCBI_TaxID=4332;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4), FUNCTION, ALTERNATIVE PROMOTER
 RP USAGE, TISSUE SPECIFICITY, DEVELOPMENTAL STAGE, AND MUTANTS E9 AND
 RP E10;
 RC STRAIN=RC11-FA;
 RA Thomas D.; Barber R.; Surdin-Kerjan Y.; Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN SEQUENCE FROM N.A.
 RA Rose M.; Koetter P.; Entian K.D./EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- SIMILARITY: Contains 2 C2H2-type zinc fingers.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sub.ch).
 CC
 DR EMBL; I26506; AAA35240; 1; -.
 DR EMBL; 249627; CA89558; 1; -.
 DR PIR; S57150; SS7150.
 DR HSSP; P07248; 1ARD.
 DR GermOnline; 141960; -.
 DR TRANSFAC; T04593; -.
 DR SGD; S0003888; M5M1.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 2.
 DR SMART; SM03355; Znf_C2H2; 2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE; PS5015; ZINC_FINGER_C2H2_2; 2.
 DR DNA-binding; Nuclear Protein; Zinc-finger; Metal-binding; Repeat.
 KW DOMAIN 64 67 POLY-PHE.

RESULT 13
 SPEN_DROME STANDARD; PRT; 5560 AA.
 ID SPEN_DROME STANDARD; PRT; 5560 AA.
 AC Q8SA83; Q9N117; Q9U6C3; Q9VPL1; Q9VP12;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Split ends protein.
 GN SPEN OR CG18497.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryote; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephdroioidea; Drosophilidae; Drosophila.
 OC NCBITaxonID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4), FUNCTION, ALTERNATIVE PROMOTER
 RP USAGE, TISSUE SPECIFICITY, DEVELOPMENTAL STAGE, AND MUTANTS E9 AND
 RP E10;
 RC D57; RT
 RC TISSUE=Embryo;
 RX MEDLINE=0025336; PubMed=10556062;
 RX MEDLINE=2015749; PubMed=10655223;
 RA McElitte E.L.; Harding K.W.; Mace K.A.; Ronshaugen M.R.; Wang F.Y.;
 RA McGinnis W.;
 RT "spen encodes an RNP motif protein that interacts with Hox Pathways
 RT trunk;"
 RT Development 126:5373-5385(1999).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 4).
 RC TISSUE=Embryo;
 RX MEDLINE=2015749; PubMed=10655223;
 RA Rebay I.; Chen P.; Hsiao P.; Kolodziej P.A.; Kuang B.H.; Laverty T.,
 RA Suh C., Voas M., Williams A., Willard G.M.;
 RA "A genetic screen for novel components of the Ras/mitogen-activated
 RT protein kinase signaling pathway that interact with the yan gene of
 RT Drosophila identifies split ends, a new RNA recognition motif-
 RT containing protein.";
 RL Genetics 154:635-712(2000).
 RL [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1), SUBCELLULAR LOCATION, AND TISSUE
 RP SPECIFICITY.
 RP TISSUE=Embryo;
 RX MEDLINE=2011275; PubMed=10704397;
 RA Kuang B.H.; Wu S.C.-Y.; Shin Y.-A.; Luo L.; Kolodziej P.A.;
 RA "split ends encode large nuclear proteins that regulate neuronal
 RT cell fate and axon extensi";
 RT Development 127:1517-1529(2000).
 RN [4]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC STRAIN=Berkeley;

- | | | |
|----|--|----|
| RT | Chen F., Rebay I.; "split ends, a new component of the <i>Drosophila</i> EGF receptor pathway, regulates development of midline glial cells.", Curr. Biol. 10:943-946 (2000). [9] | RT |
| RN | FUNCTION ON WG PATHWAY. | RN |
| RX | MEDLINE=2268876; PubMed=1278375; | RX |
| RA | Liu H.V., Dorroquez D.B., Cho S., Chen F., Rebay I., Cadigan K.M.; "Splits ends is a tissue-/promoter specific regulator of Wingless signaling". | RA |
| RT | Development 130:3125-3135 (2003). | RT |
| RL | -I- FUNCTION: Probable corepressor protein, which regulates different key pathways such as the EGF receptor and Wg pathways. Involved in neuronal cell fate, survival and axon guidance. May act with the Hox gene Deformed and the EGF receptor signalling pathway. Positive regulator of the Wg pathway in larval tissues but not in embryonic tissues. May act as a transcriptional corepressor protein, which repress transcription via the recruitment of large complexes containing histone deacetylase proteins. | RL |
| CC | -I- SUBCELLULAR LOCATION: Nucleus. | CC |
| CC | -I- ALTERNATIVE PRODUCTS: | CC |
| CC | -I- Event-Alternative promoter; Named isoforms=4; Alternative Promoters; | CC |
| CC | Event=Alternative splicing; Named isoforms=4; IsoID=Q8SX83-1; Sequence=Displayed; | CC |
| CC | Comment=2 isoforms, 1 (shown here) and 2, are produced by use of alternative promoters. | CC |
| CC | Name=1; IsoID=Q8SX83-2; Sequence=VSP_008565; VSP_008566; Note=No experimental confirmation available; | CC |
| CC | Name=3; Synonyms=SpenL; IsoID=Q8SX83-3; Sequence=VSP_008567; Note=Produced by alternative splicing of isoform 1; | CC |
| CC | Name=4; Synonyms=SpenS; IsoID=Q8SX83-4; Sequence=VSP_008565; VSP_008566; VSP_008567; Note=Produced by alternative splicing of isoform 2; | CC |
| CC | -I- TISSUE SPECIFICITY: Ubiquitous. Expressed prior to cellularization in stage 3 embryos and in blastoderm cells, including pole cells. Expressed throughout the rest of embryogenesis. Later, it is expressed at higher level in epidermal cells and CNS. | CC |
| CC | -I- DEVELOPMENTAL STAGE: Isoform 3 is expressed both maternally and zygotically. Belongs to the Spen family. | CC |
| CC | -I- SIMILARITY: Contains 3 RNA recognition motif (RRM) domains. | CC |
| CC | between the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement (See http://www.isb-sib.ch/anno.html or send an email to license@isb-sib.ch). | CC |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. | CC |
| CC | CAUTION: It is uncertain whether Met-1 or Met-7 is the initiator of the genome sequence of <i>Drosophila melanogaster</i> . This statement is not removed. Usage by commercial entities requires a license agreement (See http://www.isb-sib.ch/anno.html or send an email to license@isb-sib.ch). | CC |
| CC | REVISIONS, AND ALTERNATIVE SPLICING. | CC |
| RA | MEDLINE=22426069; PubMed=1253752; | RA |
| RA | Mira S., Crosby M.A., Campbell K.S., Matthews B.B., Campbell K.S., Hrdlicky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E., Smith C.D., Woodage T., Wasserman D.A., Weissenbach J., Wang X., Wang Z.-Y., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Williams S.M., Zaveri R., Zavestina J.S., Zhou M., Zhang G., Zhao Q., Zheng L., Ye J., Yeh R.-F., Zaveri R., Zhai J.S., Zhou X., Zhu S., Smith H.Q., Zhang X.H., Zhong F.N., Zhong W.W., Zhou X., Zhu X., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; "The genome sequence of <i>Drosophila melanogaster</i> ."; Science 287:2185-2195 (2000). | RA |
| RA | (15) | RA |
| RN | SEQUENCE OF 424-2402 FROM N.A. | RN |
| RA | STRAIN=BKTC; TISSUE=Embryo; MEDLINE=22426066; PubMed=12537569; | RA |
| RA | Smith C.D., Tupaj J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P., Bentzien B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Staglione M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.; "Annotation of the <i>Drosophila melanogaster</i> euchromatic genome: a systematic review"; Genomic Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002). | RA |
| RA | [16] | RA |
| RA | FUNCTON ON EGF RECEPTOR PATHWAY. | RN |
| RA | MEDLINE=20253107; PubMed=10790398; | RA |
| RA | Lane M.E., Elend M., Heimann D., Herr A., Marzokho S., Herzog A., George R.A., Guarin J.W., Brokstein P., Yu C., Champé M., Park S., Wan K.H., Rubin G.M.; "A Drosophila full-length cDNA resource.;" Genom Biol. 3:RESEARCH0080.1-RESEARCH0080.8 (2002). | RA |
| RA | [17] | RA |
| RP | FUNCTION ON EGF RECEPTOR PATHWAY. | RN |
| RP | MEDLINE=2414403; PubMed=1095845; | RP |
| RA | Lerner C.F.; "An screen for modifiers of cyclin E function in <i>Drosophila melanogaster</i> identifies Cdk2 mutations revealing the insignificance of putative phosphorylation sites in Cdk2."; Genetcs 155:233-244 (2000). | RA |
| RA | [18] | RA |
| RP | FUNCTION ON EGF RECEPTOR PATHWAY. | RX |
| RP | MEDLINE=20414403; PubMed=1095845; | RX |

Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

[4]	RN	SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE	CC	Name=1; IsoID=Q8VIM5-1; Sequence=Displayed;
RX	MEDLINE=22527262; PubMed=12640126;	CC	Name=2; Synonyms=BSAC2A Myocardin A; IsoID=Q8VIM5-2; Sequence=VSP_007662;	
RA	Du K.I., Ip H.S., Li J.J., Chen M., Dandre F., Yu W., Li M.M., Owens G.K., Parmacek M.S.,	CC	Note= NO EXPERIMENTAL CONFIRMATION AVAILABLE;	
RA	"Myocardin is a critical serum response factor cofactor in the transcriptional program regulating smooth muscle cell differentiation";	CC	- TISSUE SPECIFICITY: Expressed in heart, aorta, and in smooth muscle cell-containing tissues: stomach, bladder and uterus.	
RT	RT	CC	- DEVELOPMENTAL STAGE: Detected in the cardiac crescent at 7.75 dpc and in the linear heart tube at 8.0 dpc and the developing atrial and aortic ventricular chambers until birth. Also detected in a sub-vascular and visceral smooth muscle cells: aortic arch arteries at 9.5 dpc; walls of the esophagus, dorsal aorta, pulmonary outflow tract, lung, gut, stomach, small intestine, bladder, and the head mesenchymal at 13.5 dpc until birth. Not detected in skeletal muscle cells.	
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Tomaru Y., Suzuki H., Miyazawa I., Kyosawa H., Naito N., Saito R., Hasegawa Y., Negami A., Schoribach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbin L.E., Cousins S., Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson J., Jarvis E.D., Kanai A., Kawaji H., Kawashima Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lehnhard B., Lyons P.A., Magiott D.R., Maitais L., Marchlomni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavon W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultan R., Takemoto Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Waranabe Y., Wells C., Wu J., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imamura K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Saeki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Barrey E., Hayashizaki Y.,	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
RA	"Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs";	CC	- SIMILARITY: Contains 3 RPEL repeats.	
RP	SEQUENCE OF 728-935 FROM N.A.	CC	- SIMILARITY: Contains 1 SAP domain.	
RC	STRAIN=C57BL/6J; TISSUE=Embryonic heart;	CC	- SIMILARITY: Contains 1 SAP domain.	
RX	MEDLINE=22354683; PubMed=12466851;	CC	-	
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Tomaru Y., Suzuki H., Miyazawa I., Kyosawa H., Naito N., Saito R., Hasegawa Y., Negami A., Schoribach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbin L.E., Cousins S., Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson J., Jarvis E.D., Kanai A., Kawaji H., Kawashima Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lehnhard B., Lyons P.A., Magiott D.R., Maitais L., Marchlomni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavon W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultan R., Takemoto Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Waranabe Y., Wells C., Wu J., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imamura K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Saeki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Barrey E., Hayashizaki Y.,	CC	-	
RX	MEDLINE=22653656; PubMed=12663482;	CC	-	
RA	Kremer B.E., Wang D.-Z., Olson E.N., Owens G.K.; RT	CC	- FUNCTION: Transcriptional factor which uses the canonical single or multiple CARG boxes DNA sequence. Acts as a cofactor of serum response factor (SRF) with the potential to modulate SRF-target genes. Regulates the expression of a set of cardiac and smooth muscle-specific genes. Plays a crucial role in cardiogenesis and differentiation of the smooth muscle cell lineage (myogenesis). May bind matrix attachment regions through its SAP domain.	
RA	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	- SUBUNIT: Homodimer (By similarity). Interacts with SRF, its association does not depend on specific DNA sequences for ternary complex formation.	
RL	Circ. Res. 92:854-864 (2003).	CC	- SUBCELLULAR LOCATION: Nuclear, with a punctate intranuclear pattern with exclusion from nuclei.	
RX	REVIEW.	CC	- ALTERNATIVE PRODUCTS: Event=Alternative splicing; Named isoforms=2;	
RA	MEDLINE=21403251; PubMed=11511353;	CC	-	
RA	Hauschka S.D.;	CC	-	
RT	"Myocardin, a novel potentiator of SRF-mediated transcription in cardiac muscle";	CC	-	
RT	Mol. Cell 8:1-2 (2001).	CC	-	
RL	FUNCTION: Transcriptional factor which uses the canonical single or multiple CARG boxes DNA sequence. Acts as a cofactor of serum response factor (SRF) with the potential to modulate SRF-target genes. Regulates the expression of a set of cardiac and smooth muscle-specific genes. Plays a crucial role in cardiogenesis and differentiation of the smooth muscle cell lineage (myogenesis). May bind matrix attachment regions through its SAP domain.	CC	-	
CC	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	- SUBUNIT: Homodimer (By similarity). Interacts with SRF, its association does not depend on specific DNA sequences for ternary complex formation.	
CC	Circ. Res. 92:854-864 (2003).	CC	- SUBCELLULAR LOCATION: Nuclear, with a punctate intranuclear pattern with exclusion from nuclei.	
CC	Event=Alternative splicing; Named isoforms=2;	CC	-	
RP	SEQUENCE OF 728-935 FROM N.A.	CC	-	
RC	STRAIN=C57BL/6J; TISSUE=Embryonic heart;	CC	-	
RX	MEDLINE=22354683; PubMed=12466851;	CC	-	
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Tomaru Y., Suzuki H., Miyazawa I., Kyosawa H., Naito N., Saito R., Hasegawa Y., Negami A., Schoribach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbin L.E., Cousins S., Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson J., Jarvis E.D., Kanai A., Kawaji H., Kawashima Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lehnhard B., Lyons P.A., Magiott D.R., Maitais L., Marchlomni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavon W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultan R., Takemoto Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Waranabe Y., Wells C., Wu J., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imamura K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Saeki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Barrey E., Hayashizaki Y.,	CC	-	
RX	MEDLINE=22653656; PubMed=12663482;	CC	-	
RA	Kremer B.E., Wang D.-Z., Olson E.N., Owens G.K.; RT	CC	-	
RA	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	-	
RL	Circ. Res. 92:854-864 (2003).	CC	-	
RX	REVIEW.	CC	-	
RA	MEDLINE=21403251; PubMed=11511353;	CC	-	
RA	Hauschka S.D.;	CC	-	
RT	"Myocardin, a novel potentiator of SRF-mediated transcription in cardiac muscle";	CC	-	
RT	Mol. Cell 8:1-2 (2001).	CC	-	
RL	FUNCTION: Transcriptional factor which uses the canonical single or multiple CARG boxes DNA sequence. Acts as a cofactor of serum response factor (SRF) with the potential to modulate SRF-target genes. Regulates the expression of a set of cardiac and smooth muscle-specific genes. Plays a crucial role in cardiogenesis and differentiation of the smooth muscle cell lineage (myogenesis). May bind matrix attachment regions through its SAP domain.	CC	-	
CC	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	- SUBUNIT: Homodimer (By similarity). Interacts with SRF, its association does not depend on specific DNA sequences for ternary complex formation.	
CC	Circ. Res. 92:854-864 (2003).	CC	- SUBCELLULAR LOCATION: Nuclear, with a punctate intranuclear pattern with exclusion from nuclei.	
CC	Event=Alternative splicing; Named isoforms=2;	CC	-	
RP	SEQUENCE OF 728-935 FROM N.A.	CC	-	
RC	STRAIN=C57BL/6J; TISSUE=Embryonic heart;	CC	-	
RX	MEDLINE=22354683; PubMed=12466851;	CC	-	
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Tomaru Y., Suzuki H., Miyazawa I., Kyosawa H., Naito N., Saito R., Hasegawa Y., Negami A., Schoribach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbin L.E., Cousins S., Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson J., Jarvis E.D., Kanai A., Kawaji H., Kawashima Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lehnhard B., Lyons P.A., Magiott D.R., Maitais L., Marchlomni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavon W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultan R., Takemoto Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Waranabe Y., Wells C., Wu J., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imamura K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Saeki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Barrey E., Hayashizaki Y.,	CC	-	
RX	MEDLINE=22653656; PubMed=12663482;	CC	-	
RA	Kremer B.E., Wang D.-Z., Olson E.N., Owens G.K.; RT	CC	-	
RA	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	-	
RL	Circ. Res. 92:854-864 (2003).	CC	-	
RX	REVIEW.	CC	-	
RA	MEDLINE=21403251; PubMed=11511353;	CC	-	
RA	Hauschka S.D.;	CC	-	
RT	"Myocardin, a novel potentiator of SRF-mediated transcription in cardiac muscle";	CC	-	
RT	Mol. Cell 8:1-2 (2001).	CC	-	
RL	FUNCTION: Transcriptional factor which uses the canonical single or multiple CARG boxes DNA sequence. Acts as a cofactor of serum response factor (SRF) with the potential to modulate SRF-target genes. Regulates the expression of a set of cardiac and smooth muscle-specific genes. Plays a crucial role in cardiogenesis and differentiation of the smooth muscle cell lineage (myogenesis). May bind matrix attachment regions through its SAP domain.	CC	-	
CC	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	- SUBUNIT: Homodimer (By similarity). Interacts with SRF, its association does not depend on specific DNA sequences for ternary complex formation.	
CC	Circ. Res. 92:854-864 (2003).	CC	- SUBCELLULAR LOCATION: Nuclear, with a punctate intranuclear pattern with exclusion from nuclei.	
CC	Event=Alternative splicing; Named isoforms=2;	CC	-	
RP	SEQUENCE OF 728-935 FROM N.A.	CC	-	
RC	STRAIN=C57BL/6J; TISSUE=Embryonic heart;	CC	-	
RX	MEDLINE=22354683; PubMed=12466851;	CC	-	
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Tomaru Y., Suzuki H., Miyazawa I., Kyosawa H., Naito N., Saito R., Hasegawa Y., Negami A., Schoribach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbin L.E., Cousins S., Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson J., Jarvis E.D., Kanai A., Kawaji H., Kawashima Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lehnhard B., Lyons P.A., Magiott D.R., Maitais L., Marchlomni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavon W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultan R., Takemoto Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Waranabe Y., Wells C., Wu J., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imamura K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Saeki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Barrey E., Hayashizaki Y.,	CC	-	
RX	MEDLINE=22653656; PubMed=12663482;	CC	-	
RA	Kremer B.E., Wang D.-Z., Olson E.N., Owens G.K.; RT	CC	-	
RA	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	-	
RL	Circ. Res. 92:854-864 (2003).	CC	-	
RX	REVIEW.	CC	-	
RA	MEDLINE=21403251; PubMed=11511353;	CC	-	
RA	Hauschka S.D.;	CC	-	
RT	"Myocardin, a novel potentiator of SRF-mediated transcription in cardiac muscle";	CC	-	
RT	Mol. Cell 8:1-2 (2001).	CC	-	
RL	FUNCTION: Transcriptional factor which uses the canonical single or multiple CARG boxes DNA sequence. Acts as a cofactor of serum response factor (SRF) with the potential to modulate SRF-target genes. Regulates the expression of a set of cardiac and smooth muscle-specific genes. Plays a crucial role in cardiogenesis and differentiation of the smooth muscle cell lineage (myogenesis). May bind matrix attachment regions through its SAP domain.	CC	-	
CC	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	- SUBUNIT: Homodimer (By similarity). Interacts with SRF, its association does not depend on specific DNA sequences for ternary complex formation.	
CC	Circ. Res. 92:854-864 (2003).	CC	- SUBCELLULAR LOCATION: Nuclear, with a punctate intranuclear pattern with exclusion from nuclei.	
CC	Event=Alternative splicing; Named isoforms=2;	CC	-	
RP	SEQUENCE OF 728-935 FROM N.A.	CC	-	
RC	STRAIN=C57BL/6J; TISSUE=Embryonic heart;	CC	-	
RX	MEDLINE=22354683; PubMed=12466851;	CC	-	
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Tomaru Y., Suzuki H., Miyazawa I., Kyosawa H., Naito N., Saito R., Hasegawa Y., Negami A., Schoribach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbin L.E., Cousins S., Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson J., Jarvis E.D., Kanai A., Kawaji H., Kawashima Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lehnhard B., Lyons P.A., Magiott D.R., Maitais L., Marchlomni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavon W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultan R., Takemoto Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Waranabe Y., Wells C., Wu J., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imamura K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Saeki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Barrey E., Hayashizaki Y.,	CC	-	
RX	MEDLINE=22653656; PubMed=12663482;	CC	-	
RA	Kremer B.E., Wang D.-Z., Olson E.N., Owens G.K.; RT	CC	-	
RA	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	-	
RL	Circ. Res. 92:854-864 (2003).	CC	-	
RX	REVIEW.	CC	-	
RA	MEDLINE=21403251; PubMed=11511353;	CC	-	
RA	Hauschka S.D.;	CC	-	
RT	"Myocardin, a novel potentiator of SRF-mediated transcription in cardiac muscle";	CC	-	
RT	Mol. Cell 8:1-2 (2001).	CC	-	
RL	FUNCTION: Transcriptional factor which uses the canonical single or multiple CARG boxes DNA sequence. Acts as a cofactor of serum response factor (SRF) with the potential to modulate SRF-target genes. Regulates the expression of a set of cardiac and smooth muscle-specific genes. Plays a crucial role in cardiogenesis and differentiation of the smooth muscle cell lineage (myogenesis). May bind matrix attachment regions through its SAP domain.	CC	-	
CC	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	- SUBUNIT: Homodimer (By similarity). Interacts with SRF, its association does not depend on specific DNA sequences for ternary complex formation.	
CC	Circ. Res. 92:854-864 (2003).	CC	- SUBCELLULAR LOCATION: Nuclear, with a punctate intranuclear pattern with exclusion from nuclei.	
CC	Event=Alternative splicing; Named isoforms=2;	CC	-	
RP	SEQUENCE OF 728-935 FROM N.A.	CC	-	
RC	STRAIN=C57BL/6J; TISSUE=Embryonic heart;	CC	-	
RX	MEDLINE=22354683; PubMed=12466851;	CC	-	
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Tomaru Y., Suzuki H., Miyazawa I., Kyosawa H., Naito N., Saito R., Hasegawa Y., Negami A., Schoribach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbin L.E., Cousins S., Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson J., Jarvis E.D., Kanai A., Kawaji H., Kawashima Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lehnhard B., Lyons P.A., Magiott D.R., Maitais L., Marchlomni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavon W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultan R., Takemoto Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Waranabe Y., Wells C., Wu J., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imamura K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Saeki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Barrey E., Hayashizaki Y.,	CC	-	
RX	MEDLINE=22653656; PubMed=12663482;	CC	-	
RA	Kremer B.E., Wang D.-Z., Olson E.N., Owens G.K.; RT	CC	-	
RA	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	-	
RL	Circ. Res. 92:854-864 (2003).	CC	-	
RX	REVIEW.	CC	-	
RA	MEDLINE=21403251; PubMed=11511353;	CC	-	
RA	Hauschka S.D.;	CC	-	
RT	"Myocardin, a novel potentiator of SRF-mediated transcription in cardiac muscle";	CC	-	
RT	Mol. Cell 8:1-2 (2001).	CC	-	
RL	FUNCTION: Transcriptional factor which uses the canonical single or multiple CARG boxes DNA sequence. Acts as a cofactor of serum response factor (SRF) with the potential to modulate SRF-target genes. Regulates the expression of a set of cardiac and smooth muscle-specific genes. Plays a crucial role in cardiogenesis and differentiation of the smooth muscle cell lineage (myogenesis). May bind matrix attachment regions through its SAP domain.	CC	-	
CC	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	- SUBUNIT: Homodimer (By similarity). Interacts with SRF, its association does not depend on specific DNA sequences for ternary complex formation.	
CC	Circ. Res. 92:854-864 (2003).	CC	- SUBCELLULAR LOCATION: Nuclear, with a punctate intranuclear pattern with exclusion from nuclei.	
CC	Event=Alternative splicing; Named isoforms=2;	CC	-	
RP	SEQUENCE OF 728-935 FROM N.A.	CC	-	
RC	STRAIN=C57BL/6J; TISSUE=Embryonic heart;	CC	-	
RX	MEDLINE=22354683; PubMed=12466851;	CC	-	
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Tomaru Y., Suzuki H., Miyazawa I., Kyosawa H., Naito N., Saito R., Hasegawa Y., Negami A., Schoribach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbin L.E., Cousins S., Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson J., Jarvis E.D., Kanai A., Kawaji H., Kawashima Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lehnhard B., Lyons P.A., Magiott D.R., Maitais L., Marchlomni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavon W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultan R., Takemoto Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Waranabe Y., Wells C., Wu J., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imamura K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Saeki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Barrey E., Hayashizaki Y.,	CC	-	
RX	MEDLINE=22653656; PubMed=12663482;	CC	-	
RA	Kremer B.E., Wang D.-Z., Olson E.N., Owens G.K.; RT	CC	-	
RA	"Myocardin is a key regulator of CAG-dependent transcription of multiple smooth muscle marker genes";	CC	-	
RL	Circ. Res. 92:854-864 (2003).	CC	-	
RX	REVIEW.	CC	-	
RA	MEDLINE=21403251; PubMed=11511353;	CC	-	
RA	Hauschka S.D.;	CC	-	
RT	"Myocardin, a novel potentiator of SRF-mediated transcription in cardiac muscle";	CC	-	
RT	Mol. Cell 8:1-2 (2001).	CC	-	
RL	FUNCTION: Transcriptional factor which uses the canonical single or multiple CARG boxes DNA sequence. Acts as a cofactor of serum response factor (SRF) with the potential to modulate SRF-target genes. Regulates the expression of a set of cardiac and smooth muscle-specific genes. Plays a crucial role in cardiogenesis and differentiation of the smooth muscle cell lineage			

FT CONFLICT 110 110 E -> EA (IN REF. 4).
FT CONFLICT 794 794 G -> D (IN REF. 3 AND 5).
SQ SEQUENCE 935 AA; 101373 MW; C1DB83FFBC181149 CRC64;

Query Match 93.1%; Score 27; DB 1; Length 935;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SSGPSL 6
Db 906 SSGPSI 911

Search completed. March 10, 2004, 09:13:52
Job time : 3.09728 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 1.8677 Seconds
 (without alignments)
 309.015 Million cell updates/sec

Title: US-09-848-834A-6

Perfect score: 29

Sequence: 1 SSGPSL 6

Scoring table: BloSUM62
 Gapop 10.0 , Gapext 0.5
 Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing First 45 summaries

Database : PIR78:*

1: Pir1:*

2: piz2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	29	100.0	225	1	JQ2010	nonstructural protein V - simian paramyxovirus SV41
2	29	100.0	280	2	A5307	C;Species: Simian paramyxovirus SV41
3	29	100.0	348	2	AB3260	C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Jul-1999
4	29	100.0	394	2	JQ2041	C;Accession: JQ0040
5	29	100.0	405	2	B97171	R;Kawano, M.; Tsurudome, M.; Oki, N.; Nishio, M.; Komada, H.; Kusagawa, J. Gen. Virol. 74, 911-916, 1993
6	29	100.0	532	2	S7453	A;Title: Sequence determination of the P gene of simian virus 41: presence of irregular
7	29	100.0	989	2	T46659	A;Reference number: JQ2040; PMID:53260108; PMID:8492098
8	29	100.0	997	2	S44457	A;Accession: JQ2040
9	29	100.0	1001	2	S30285	A;Molecule type: Genomic RNA
10	29	100.0	1074	2	T04777	A;Residues: 1-225 <KRW>
11	29	100.0	1380	2	S57150	A;Cross-references: GB:S60811; NID:9385516; PIDN:AAB26639_1; PID:9385517
12	27	93.1	222	2	T34050	C;Keywords: alternative splicing; nonstructural protein V; V/P protein homology <P/N>
13	27	93.1	1487	2	T02880	F;-163;Domain: V/P Protein homology <P/N>
14	27	93.1	1564	2	T27121	C;Genetics:
15	26	89.7	52	2	C38230	A;Gene: V
16	26	89.7	76	2	C36817	A;Status: V
17	26	89.7	102	1	CCEG	A;Accession: A53027
18	26	89.7	102	2	R83089	N;Alternative names: hairy protein homolog
19	26	89.7	137	2	G75471	C;Accession: A53027
20	26	89.7	156	2	H72621	R;Feder, J.N.; Li, L.; Jan, Y.N.
21	26	89.7	209	1	HSXL1A	Genomics 20, 56-61, 1994
22	26	89.7	229	2	I51227	A;Title: Genomic cloning and chromosomal localization of HRY, the human homolog to the
23	26	89.7	260	2	AH2878	A;Reference number: A53027; PMID:94292187; PMID:8020957
24	26	89.7	260	2	B97655	A;Gene: GDB:HRY
25	26	89.7	263	2	A87149	A;Cross-references: GDB:217077; OMIM:139605
26	26	89.7	285	2	D70566	A;Map Position: 3q28-3q29
27	26	89.7	297	2	T21387	C;Keywords: immediate-early protein
28	26	89.7	302	2	G70614	Query Match 100.0%; Score 29; DB 2; Length 280;
29	26	89.7	326	2	JQ1659	Query Match 100.0%; Score 29; DB 2; Length 280;

RESULT 1

JQ0040

demonstrational protein V - simian paramyxovirus SV41

C;Species: Simian paramyxovirus SV41

C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Jul-1999

C;Accession: JQ0040

R;Kawano, M.; Tsurudome, M.; Oki, N.; Nishio, M.; Komada, H.; Kusagawa, J. Gen. Virol. 74, 911-916, 1993

A;Title: Sequence determination of the P gene of simian virus 41: presence of irregular

A;Reference number: JQ2040; PMID:53260108; PMID:8492098

A;Accession: JQ2040

A;Molecule type: Genomic RNA

A;Residues: 1-225 <KRW>

A;Cross-references: GB:S60811; NID:9385516; PIDN:AAB26639_1; PID:9385517

C;Keywords: alternative splicing; nonstructural protein V; V/P protein homology <P/N>

F;-163;Domain: V/P Protein homology <P/N>

P;-163;Domain: V/P Protein homology <P/N>

C;Genetics:

A;Gene: V

A;Status: V

A;Accession: A53027

N;Alternative names: hairy protein homolog

C;Accession: A53027

R;Feder, J.N.; Li, L.; Jan, Y.N.

Genomics 20, 56-61, 1994

A;Title: Genomic cloning and chromosomal localization of HRY, the human homolog to the

A;Reference number: A53027; PMID:94292187; PMID:8020957

A;Gene: GDB:HRY

A;Cross-references: GDB:217077; OMIM:139605

A;Map Position: 3q28-3q29

C;Keywords: immediate-early protein

Query Match 100.0%; Score 29; DB 1; Length 225;

Best Local Similarity 100.0%; Pred. NO. 68;

Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

Query 1 SSGPSL 6

Db 144 SSGPSL 149

RESULT 2

A53027

transcription factor HES-1 - human

C;Species: Homo sapiens (man)

C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 03-Dec-1999

C;Accession: A53027

R;Feder, J.N.; Li, L.; Jan, Y.N.

Genomics 20, 56-61, 1994

A;Title: Genomic cloning and chromosomal localization of HRY, the human homolog to the

A;Reference number: A53027; PMID:94292187; PMID:8020957

A;Gene: GDB:HRY

A;Cross-references: GDB:217077; OMIM:139605

A;Map Position: 3q28-3q29

C;Keywords: immediate-early protein

Query Match 100.0%; Score 29; DB 2; Length 280;

Best Local Similarity 100.0%; Pred. NO. 68;

Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

Query 1 SSGPSL 6

Db 144 SSGPSL 149

Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
Db 264 SSGPSL 269

RESULT 3
AB3260 hypothetical membrane spanning protein BMEI0063 [imported] - Brucella melitensis (strain C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AB3260
R:Delvecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N.; Mazur, M.; Gotsman, E.; Selkov, E.; Blazquez, P.H.; Hagiwara, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 43-48, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AB3252; PMID:11756688
A:Accession: AB3260
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <KRB>
A:Cross-references: GB:AE008917; PIDN:AA51245.1; PID:917981935; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0063
A:Map Position: 1

Query Match 100.0%; Score 29; DB 2; Length 348;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
Db 140 SSGPSL 145

RESULT 4
JQ2041 polymerase-associated nucleocapsid phosphoprotein - simian paramyxovirus SV41
C:Species: simian paramyxovirus SV41
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 16-Jul-1999
C:Accession: JQ2041
R:Kawano, M.; Tsurudome, M.; Oki, N.; NishiO, M.; Komada, H.; Matsumura, H.; Kusagawa, S.; J. Gen. Virol. 74, 911-916, 1993
A:Title: Sequence determination of the P gene of simian virus 41: presence of irregular A:Reference number: JQ2040; MUID:93260408; PMID:8492098
A:Accession: JQ2041
A:Molecule type: RNA
A:Residues: 1-394 <KAW>
C:Genetics:

A:Gene: P
C:Superfamily: simian paramyxovirus P protein; V/P protein homology
C:Keywords: nucleocapsid; phosphoprotein; RNA editing
P:1-163/Domain: V/P protein homology <VPN>

Query Match 100.0%; Score 29; DB 2; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
Db 144 SSGPSL 149

RESULT 5
B97171 uncharacterized conserved protein CAC2200 [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: B97171

Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
Db 198 SSGPSL 203

RESULT 6
S74453 hypothetical protein Slr1484 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
A:Cross-references: EMBL:S74453; PIDN:91650; PID:BA16605.1; PID:d10173
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O., K.; Okumura, S.; Shimpou, S.; Takeuchi, C.; Wada, T.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74453
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-532 <KAN>
A:Cross-references: EMBL:D90899; GB:AB001339; PID:gi1650; PIDN:91650; PID:BA16605.1; PID:d10173
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
A:Cross-references: EMBL:T46659; PIDN:91653
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 01-Dec-2000
R:Aronson, B.D.; Johnson, K.A.; Dunlap, J.C.; Aronson, B.D.; Johnson, K.A.; Dunlap, J.C.; Proc. Natl. Acad. Sci. U.S.A. 91, 7683-7687, 1994
A:Title: Circadian clock locus frequency: Protein encoded by a single open reading frame A:Reference number: Z23120; MUID:94329580; PMID:8052643
A:Accession: T46659
A:Status: preliminary; translated from GB/EMBL/DDJB
A:Molecule type: mRNA
A:Residues: 1-989 <ZAR>
A:Cross-references: EMBL:U17073; PIDN:955951; PID:AAA57121.1; PID:9595952
A:Experimental source: strain OR 74 A; isolate bda; mycelia
R:McClung, C.R.; Cox, B.A.; Dunlap, J.C.
Nature 339, 558-562, 1989
A:Title: The Neurospora circadian clock gene frequency shares a sequence element with the Drosophila melanogaster circadian clock gene
A:Reference number: 804653; MUID:98281721; PMID:2525233

A;Accession: S04653
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 'NCNN' 206, 'P', 208-232, 'H', 234-989 <NCC>
A;Note: the authors translated the initiation codon GTG as Val
C;Gene:
A;GeneID: frq
A;Map position: VII R

Query Match 100.0%; Score 29; DB 2; Length 999;
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 1 SSGPSL 6
Db 271 SSGPSL 276

RESULT 8
S44457
Period clock protein frq - *Sordaria fimicola*
C;Species: *Sordaria fimicola*
C;Accession: S44457
C;Date: 29-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 05-Nov-1999
R;Merrow, M.W.; Dunlap, J.C.
EMBO J. 13: 2257-2266, 1994
A;Title: Intergeneric complementation of a circadian rhythmicity defect: phylogenetic co
A;Reference number: S44457; PMID:8194516
A;Accession: S44457
A;Molecule type: DNA
A;Residues: 1-997 <MBR>
A;Cross-references: EMBL:L14467; NID:9310366; PID:9530050
C;Genetics:

C;Keywords: circadian rhythm

Query Match 100.0%; Score 29; DB 2; Length 997;
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 1 SSGPSL 6
Db 277 SSGPSL 282

RESULT 9
S0385
G9a Protein - human
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C;Accession: S30385
R;Milner, C.M.; Campbell, R.D.
Biochem. J. 290, 811-818, 1993
A;Title: The G9a gene in the human major histocompatibility complex encodes a novel prob
A;Reference number: S30385; PMID:8457211
A;Accession: S30385
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1001 <MLI>
A;Cross-references: EMBL:X69838; NID:9287864; PID:CAA49491_1; PID:9287865
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F;641-673/Domain: ankyrin repeat homology <ANR>

Query Match 100.0%; Score 29; DB 2; Length 1001;
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 1 SSGPSL 6
Db 396 SSGPSL 401

RESULT 10
T34050
hypothetical protein P28B3.2 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 04-Mar-2000
C;Accession: T34050
R;Geissel, C.; Kramer, J.; Smith, A.
Submitted to the EMBL Data Library, May 1997
A;Description: The sequence of C. elegans cosmid F28B3.
A;Reference number: 221469

A;Accession: T34060
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Cross-references: EMBL:AF0031136; PIDN:AAB93634.1; GSPDB:GN00019; CESP:F28B3.2
A;Experimental source: strain Bristol N2; Clone F28B3
C;Genetics:
A;Map position: 1
A;Superfamily: Caenorhabditis elegans hypothetical protein F28B3.2
C;Species: Caenorhabditis elegans hypothetical protein F28B3.2
Query Match Score 27; DB 2; Length 222;
Best Local Similarity 93.1%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SSGPSL 6
Db 28 SSGPSI 33

RESULT 13
T02550 hypothetical protein L1439.2 [imported] - Leishmania major (strain Friedlin)
C;Species: Leishmania major
C;Date: 24-Mar-1999 #text_change 19-May-2000
C;Accession: F81462; T02550
R;Myler, P.J.; Audlemann, L.; devos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.; Proc. Natl. Acad. Sci. U.S.A. 96, 2903-2908, 1999
A;Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-coding genes
A;Reference number: A81455; MUID:99178987; PMID:10077609
A;Accession: F81462
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1487 <PYL>
A;Cross-references: GB:AB001274; NID:33264550; PIDN: AAC24673.1; PID:G2266918; GSPDB:GN00019
A;Experimental source: strain MHOM/IL/81/Friedlin
C;Genetics:
A;Gene: L1439.2
A;Map Position: 1

Query Match Score 93.1%; DB 2; Length 1487;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SSGPSL 6
Db 1478 SSGPSI 1482

RESULT 14
T27121 hypothetical protein Y53C10A.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T27121
P;White, S.
submitted to the EMBL Data Library, November 1998
A;Reference number: 220314
A;Accession: T27121
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1564 <WTL>
A;Cross-references: EMBL:AL033536; PIDN:CAA22142.1; CBSP:Y53C10A.9
A;Experimental source: clone Y53C10A
C;Genetics:
A;Gene: CESP:Y53C10A.9
A;Introns: 43/3; 92/2; 148/2; 226/3; 354/1; 712/3; 817/1; 875/1; 916/3; 984/3; 1059/2;
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
Query Match Score 93.1%; DB 2; Length 1564;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Sequence 39, App1									
16	21	100.0	14	14	US-10-083-768-39				
17	21	100.0	14	14	US-10-083-768-216				
18	21	100.0	15	9	US-09-953-510-67				
19	21	100.0	15	9	US-09-953-510-68				
20	21	100.0	15	11	US-09-953-413-67				
21	21	100.0	15	11	US-09-953-413-68				
22	21	100.0	15	14	US-10-117-255-67				
23	21	100.0	15	14	US-10-147-255-68				
24	21	100.0	18	14	US-10-225-567A-1363				
25	21	100.0	20	9	US-09-884-761-3-3705				
26	21	100.0	20	9	US-09-884-761-3-6310				
27	21	100.0	20	10	US-10-962-756-845				
28	21	100.0	20	15	US-10-253-471-845				
29	21	100.0	20	16	US-10-53-492-845				
30	21	100.0	21	13	US-10-153-064-74				
31	21	100.0	21	14	US-10-153-604A-74				
32	21	100.0	21	14	US-10-153-258				
33	21	100.0	26	14	US-10-208-304-5				
34	21	100.0	28	9	US-09-848-834A-11				
35	21	100.0	29	9	US-09-864-761-46228				
36	21	100.0	30	10	US-09-974-879-334				
37	21	100.0	30	10	US-09-563-222-154				
38	21	100.0	30	10	US-09-736-334				
39	21	100.0	30	11	US-09-848-683-334				
40	21	100.0	31	9	US-09-848-834A-9				
41	21	100.0	31	9	US-09-848-834A-15				
42	21	100.0	31	14	US-10-029-386-30958				
43	21	100.0	33	9	US-09-848-834A-12				
44	21	100.0	33	14	US-10-012-957A-213				
45	21	100.0	34	9	US-09-848-834A-10				
ALIGNMENTS									
RESULT 1									
US-09-848-834A-5									
Sequence 5, Application US/09848834A									
; Patent No. US20020076416A1									
; GENERAL INFORMATION:									
; TITLE OF INVENTION: Chimeric Peptide Immunogens									
; FILE REFERENCE: 11-02865-0047									
; CURRENT APPLICATION NUMBER: US/09-848-834A									
; CURRENT FILING DATE: 2001-05-04									
; PRIORITY APPLICATION NUMBER: 60/202,328									
; PRIORITY FILING DATE: 2000-05-05									
; NUMBER OF SEQ ID NOS: 20									
; SOFTWARE: PatentIn version 3.0									
SEQ ID NO 5									
LENGTH: 4									
TYPE: PRT									
ORGANISM: Artificial Sequence									
FEATURE:									
OTHER INFORMATION: Synthetic peptide									
US-09-848-834A-5									
Query Match 100.0% Score 21; DB 9; Length 4;									
Best Local Similarity 100.0% Pred. No. 7.1e+05; Mismatches 0; Indels 0; Gaps 0;									
Qy 1 GPSL 4									
Db 1 GPSL 4									
RESULT 2									
US-10-285-976-224									
Sequence 224, Application US/10285976									
; Sequence 224, Application US/10285976									
; Publication No. US20030165500A1									
; GENERAL INFORMATION:									
; APPLICANT: Rhee, Chae-Seo									
; APPLICANT: Malini, Sen									
SUMMARIES									
%									
Description									
ID DB Length Match Score No. Result									
1	21	100.0	4	9	US-09-848-834A-5	Sequence 5, Appl			
2	21	100.0	4	14	US-10-295-975-224	Sequence 24, App			
3	21	100.0	6	9	US-09-848-834A-6	Sequence 6, Appl			
4	21	100.0	6	14	US-10-223-711-4	Sequence 4, Appl			
5	21	100.0	8	9	US-09-848-834A-7	Sequence 7, Appl			
6	21	100.0	9	14	US-10-102-622-2	Sequence 2, Appl			
7	21	100.0	10	15	US-10-371-069-209	Sequence 209, APP			
8	21	100.0	10	15	US-10-371-645-209	Sequence 209, APP			
9	21	100.0	10	15	US-10-371-260-209	Sequence 209, APP			
10	21	100.0	10	15	US-10-373-238-49	Sequence 49, Appl			
11	21	100.0	11	15	US-10-371-069-224	Sequence 224, APP			
12	21	100.0	11	15	US-10-371-645-224	Sequence 224, APP			
13	21	100.0	11	15	US-10-371-260-224	Sequence 224, APP			
14	21	100.0	12	15	US-09-954-349-349	Sequence 149, APP			
15	21	100.0	12	15	US-10-024-123-10	Sequence 10, APP			

APPLICANT: Wu, Christina
 APPLICANT: Leoni, Lorenzo M.
 APPLICANT: Corr, Maripat
 APPLICANT: Carron, Dennis A.
 APPLICANT: The Regents of the University of California
 TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
 FILE REFERENCE: 023070-13032US
 CURRENT APPLICATION NUMBER: US/10/285,976
 CURRENT FILING DATE: 2002-11-01
 PRIORITY NUMBER: 60/287,995
 PRIORITY FILING DATE: 2001-05-01
 PRIORITY APPLICATION NUMBER: WO PCT/US02/13602
 PRIORITY FILING DATE: 2002-05-01
 NUMBER OF SEQ ID NOS: 232
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 224
 LENGTH: 4
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURES:
 OTHER INFORMATION: Description of Artificial Sequence:short linker
 US-10-285,976-224

Query Match 100.0%; Score 21; DB 14; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
 US-09-848-834A-7
 Sequence 7, Application US/09848834A
 ; Patent No. US20020076416A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abhton Corporation
 ; TITLE OF INVENTION: Chimeric Peptide Immunogens
 ; FILE REFERENCE: 1102865-0017
 ; CURRENT APPLICATION NUMBER: US/09/848,834A
 ; CURRENT FILING DATE: 2001-05-03
 ; PRIORITY NUMBER: 60/202,328
 ; PRIOR FILING DATE: 2000-05-05
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO: 7
 ; LENGTH: 8
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic peptide
 US-09-848-834A-7

Query Match 100.0%; Score 21; DB 9; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
 US-10-102-622-2
 Sequence 2, Application US/10102622
 ; Publication No. US20030190308A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Braun, Ralph P.
 ; APPLICANT: Thomassen, Lindy
 ; APPLICANT: Van-Wely, Catherine
 ; APPLICANT: Ertl, Peter
 ; TITLE OF INVENTION: Adjuvant
 ; FILE REFERENCE: 031267-015
 ; CURRENT APPLICATION NUMBER: US/10/102,622
 ; CURRENT FILING DATE: 2002-03-19
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 2

Query Match 100.0%; Score 21; DB 9; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
 US-10-223-711-4
 Sequence 4, Application US/10223711
 ; Publication No. US2003011344A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bakalatz, Lauren O.
 ; APPLICANT: Kaufman, Pravin T.P.
 ; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides

```

; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HSV CD8 in BALB/C mice
US-10-102-622-2

Query Match          100.0%;  Score 21;  DB 14;  Length 9;
Best Local Similarity      100.0%;  Pred. No. 7.1e+05;  Indels 0;  Gaps 0;
Matches 4;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
Qy   1 GPSL 4
Db    2 GPSL 5

RESULT 7
US-10-371-069-209
; Sequence 209, Application US/10371069
; Publication No. US20030216342A1
; GENERAL INFORMATION;
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Cheanut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.10
; CURRENT APPLICATION NUMBER: US/10/371,069
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 00/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 209
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HBV 360 (peptide 1039.01)
US-10-371-069-209

Query Match          100.0%;  Score 21;  DB 15;  Length 10;
Best Local Similarity      100.0%;  Pred. No. 3.7e+02;  Indels 0;  Gaps 0;
Matches 4;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
Qy   1 GPSL 4
Db    2 GPSL 5

RESULT 8
US-10-371-645-209
; Sequence 209, Application US/10371645
; Publication No. US20030216343A1
; GENERAL INFORMATION;
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Cheanut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.11
; CURRENT APPLICATION NUMBER: US/10/371,260
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 00/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 209
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HBV 360 (peptide 1039.01)
US-10-371-645-209

Query Match          100.0%;  Score 21;  DB 15;  Length 10;
Best Local Similarity      100.0%;  Pred. No. 3.7e+02;  Indels 0;  Gaps 0;
Matches 4;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
Qy   1 GPSL 4
Db    2 GPSL 5

RESULT 9
US-10-371-260-209
; Sequence 209, Application US/10371260
; Publication No. US20030220285A1
; GENERAL INFORMATION;
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Cheanut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.13
; CURRENT APPLICATION NUMBER: US/10/371,260
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 00/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 209
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HBV 360 (peptide 1039.01)
US-10-371-260-209

Query Match          100.0%;  Score 21;  DB 15;  Length 10;
Best Local Similarity      100.0%;  Pred. No. 3.7e+02;  Indels 0;  Gaps 0;
Matches 4;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
Qy   1 GPSL 4
Db    2 GPSL 5

RESULT 10
US-10-373-238-49
; Sequence 49, Application US/10373238
; Publication No. US2004014083A1
; GENERAL INFORMATION;
; APPLICANT: General Atomics
; TITLE OF INVENTION: Detection of Heteroduplex Polynucleotides Using Mutant
; FILE REFERENCE: 39963-20022.11

```

; TITLE OF INVENTION: NUCLEIC ACID REPAIR ENZYMES WITH ATTENUATED CATALYTIC ACTIVITY
 ; FILE REFERENCE: 46699-20004.20
 ; CURRENT APPLICATION NUMBER: US/10/373,218
 ; CURRENT FILING DATE: 2003-02-24
 ; PRIORITY FILING DATE: 2000-02-14, 016
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 49
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Pb (III) ion
 ; OTHER INFORMATION: binding protein sequence.
 ; PUBLICATION INFORMATION:
 ; PATENT DOCUMENT NUMBER: S_679_548
 ; PATENT FILING DATE: 1993-06-14
 ; PRIORITY APPLICATION NUMBER: US/10-373-238-49
 ; PUBLICATION DATE: 1997-10-21
 ; US-10-373-238-49

Query Match 100.0%; Score 21; DB 15; Length 10;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 GPSL 4
D _b	2 GPSL 5

RESULT 11
 US-10-371-069-224
 ; Sequence 224, Application US/10371069
 ; Publication No. US2003021632A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EPIMMUNE Inc.
 ; APPLICANT: Fikes, John D.
 ; APPLICANT: Sette, Alessandro
 ; APPLICANT: Ishioka, Glenn Y.
 ; APPLICANT: Livingston, Brian
 ; APPLICANT: Chesnut, Robert W.
 ; APPLICANT: Epimmune Inc.
 ; TITLE OF INVENTION: Expression Vectors for Stimulating an Immune Response and Methods of Using the Same
 ; FILE REFERENCE: 39663-20022.10
 ; CURRENT APPLICATION NUMBER: US/10/371,069
 ; CURRENT FILING DATE: 2003-02-21
 ; PRIORITY FILING DATE: 1998-05-13
 ; PRIORITY APPLICATION NUMBER: US 60/085,751
 ; PRIORITY FILING DATE: 1998-05-15
 ; NUMBER OF SEQ ID NOS: 463
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 224
 ; LENGTH: 11
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: HBV env 359 (peptide 1039.06)
 ; US-10-371-069-224

Query Match 100.0%; Score 21; DB 15; Length 11;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 GPSL 4
D _b	7 GPSL 10

RESULT 12
 US-10-371-645-224
 ; Sequence 224, Application US/10371260
 ; Publication No. US2003020285A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EPIMMUNE Inc.
 ; APPLICANT: Fikes, John D.
 ; APPLICANT: Sette, Alessandro
 ; APPLICANT: Ishioka, Glenn Y.
 ; APPLICANT: Livingston, Brian
 ; APPLICANT: Chesnut, Robert W.
 ; APPLICANT: Epimmune Inc.
 ; TITLE OF INVENTION: Expression Vectors for Stimulating an Immune Response and Methods of Using the Same
 ; FILE REFERENCE: 39663-20022.13
 ; CURRENT APPLICATION NUMBER: US/10/371,260
 ; CURRENT FILING DATE: 2003-02-21
 ; PRIORITY FILING DATE: 1998-05-13
 ; PRIORITY APPLICATION NUMBER: US 60/085,751
 ; PRIORITY FILING DATE: 1998-05-15
 ; NUMBER OF SEQ ID NOS: 463
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 224
 ; LENGTH: 11
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: HBV env 359 (peptide 1039.06)
 ; US-10-371-645-224

Query Match 100.0%; Score 21; DB 15; Length 11;
 Best Local Similarity 100.0%; Pred. No. 4e+02;

	Matches	4;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	Qy	1 GPSL 4	Db	7 GPSL 10
--	---------	----	--------------	----	------------	----	--------	----	------	----	----	----------	----	-----------

RESULT 14
US-09-954-385-349
; Sequence 349, Application US/0954385
; Publication No. US20030100467A1
GENERAL INFORMATION:
; APPLICANT: Aehle, Wolfgang
; APPLICANT: Baldwin, Toby L.
; APPLICANT: Van Gastel, Franciscus J.C.
; APPLICANT: Janssen, Gisele G.
; APPLICANT: Murray, Christopher J.
; APPLICANT: Wang, Huaming
; APPLICANT: Winezyk, Deborah S.
TITLE OF INVENTION: Binding Phenol Oxidizing Enzyme-peptide
TITLE OF INVENTION: Complexes
FILE REFERENCE: GC690
CURRENT APPLICATION NUMBER: US/09/954,385
CURRENT FILING DATE: 2001-09-12
NUMBER OF SEQ ID NOS: 433
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 349
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: binding peptide

US-09-954-385-349
Query Match 100.0%; Score 21; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPSL 4
Db 8 GPSL 11

RESULT 15
US-10-024-123-10
; Sequence 10, Application US/10024123
; Publication No. US0030022263A1
GENERAL INFORMATION:
; APPLICANT: Kastan, Michael
; APPLICANT: Cannan, Christine
; APPLICANT: Kim, Seong-Tae
; APPLICANT: Lim, Dae-Sik
; APPLICANT: St. Jude Children's Research Hospital
TITLE OF INVENTION: ATM Kinase Modulation for Screening and
TITLE OF INVENTION: Therapies
FILE REFERENCE: 2427/1P142
CURRENT APPLICATION NUMBER: US/10/024,123
CURRENT FILING DATE: 2001-12-17
PRIORITY NUMBER: 03/400,653
PRIORITY FILING DATE: 1999-09-21
PRIORITY APPLICATION NUMBER: 09/248,061
PRIORITY FILING DATE: 1999-02-10
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO: 10
LENGTH: 14
TYPE: PRT
ORGANISM: Homo sapiens

US-10-024-123-10
Query Match 100.0%; Score 21; DB 14; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: March 10, 2004, 10:25:47
Job time : 3.14397 secs

Gencore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - Protein search, using SW mode

Run on: March 10, 2004, 08:58:53 ; Search time 0.731517 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834a-5

Perfect score: 21

Sequence: 1 GSSL 4

Scoring table: BLOSUM62

GapOp 10.0 , Gapext 0.5

Searched: 141681 seqs, 5207155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42_*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	46	1 DIUH_DIPPU	DIUH diploptera
2	21	100.0	46	1 DIUH_LOCMI	DIUH_locusta
3	21	100.0	46	1 DIUH_PERAM	DIUH_periplaneta
4	21	100.0	59	1 NUSM_ARTSA	NUSM_artemia
5	21	100.0	79	1 MAUD_METFL	MAUD_methiobacillus
6	21	100.0	83	1 NI9M_BOVIN	NI9M_bos
7	21	100.0	84	1 NI9M_HUMAN	NI9M_homo
8	21	100.0	84	1 NI9M_MOUSE	NI9M_mus
9	21	100.0	96	1 CY22_RHOFU	CY22_rhoospirilus
10	21	100.0	96	1 PRGB_HUMAN	PRGB_homo
11	21	100.0	96	1 RS6_MCLB	RS6_mycobacterium
12	21	100.0	97	1 CY22_RHOMO	CY22_rhodospirilus
13	21	100.0	99	1 CY21_RHOFU	CY21_rhodospirilus
14	21	100.0	100	1 CY21_RHOMO	CY21_rhodospirilus
15	21	100.0	102	1 CYC_EGRGR	CYC_euglena
16	21	100.0	105	1 C55L_SYNY3	C55L_synthochloris
17	21	100.0	106	1 CY2_EHOGI	CY2_rhodospirilus
18	21	100.0	107	1 CY2_RHOAC	CY2_rhodospirilus
19	21	100.0	108	1 NOLE_RHILP	NOLE_rhizobium
20	21	100.0	110	1 HV47_HOUSE	HV47_mus
21	21	100.0	113	1 TCLA_HUMAN	TCLA_homo
22	21	100.0	114	1 NEUL_HUMAN	NEUL_homo
23	21	100.0	117	1 RL22_SYNP6	RL22_synechococcus
24	21	100.0	120	1 C55O_BACSU	C55O_bacillus
25	21	100.0	121	1 RL22_SYNY3	RL22_synechocystis
26	21	100.0	122	1 Y764_XANAC	Y764_xanthomonas
27	21	100.0	125	1 KC1A_PIG	KC1A_scarota
28	21	100.0	125	1 NEUL_HUMAN	NEUL_homo
29	21	100.0	125	1 NEUL_TIG	NEUL_sus_scrofa
30	21	100.0	134	1 BPPI_MOUSE	BPPI_mus_musculus
31	21	100.0	138	1 BGAL_AGRRD	BGAL_agrobacterium
32	21	100.0	139	1 A55A_MYCMR	A55A_mycobacterium
33	21	100.0	140	1 CENA_HUMAN	CENA_homo_sapiens

RESULT 1
DIUH_DIPPU STANDARD: PRT; 46 AA.
ID DIUH_DIPPU
AC P83273;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Diuretic hormone Class I (Diuretic peptide) (DP) (DH(46)).
OS Diptoptera punctata (Pacific beetle cockroach).
OC Eubaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC Blaberidae; Diptoptera.
NCI_TaxID=694;
RN [1];
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE=brain, and CARICA caricae;
MEDLINE=2030924; PubMed=10841553;
RA Furuya K., Milchak R.J., Schegg K.M., Zhang J., Tobe S.S., Coast G.M.,
Scholle D.A.;
RA "Cockroach diuretic hormones: characterization of a calcitonin-like
peptide in insects";
RL Proc. Natl. Acad. Sci. U.S.A. 97:6469-6474 (2000);
CC -1- FUNCTION: Regulation of fluid secretion. Stimulates primary urine
secretion by Malpighian tubules and causes a dose-dependent effect on
the transport of CAMP levels in the tubules. Has a greater effect on
co-occurs with it.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MASS SPECTROMETRY: MW=5322; MW_ERR=0.1; METHOD=Electrospray.
CC -1- SIMILARITY: Belongs to the savagine/corticotropin-releasing
factor/urotensin I family.
CC InterPro_IPR00087; Corticotropin.
DR InterPro_IPR00361; Diuretic_hormn.
DR Pfam_Pf00473; CRP; 1.
DR ProDom_PDO14750; Diurtc_hormn; 1.
DR SMART_SM0009; CRP; 1.
DR PROSITE_PS0511; CRF; 1.
KW Hormone; Amidation.
FT MOD_RES 46 AMIDATION.
SQ SEQUENCE 46 AA; 5322 MW; 693CD5A16E47F67E CRC64;

Query Match Score 21; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPSI 4
Db 4 GPSI 7

RESULT 2
DIUH_LOCMI STANDARD: PRT; 46 AA.
ID DIUH_LOCMI
AC P2465;
DT 01-NOV-1991 (Rel. 20, Created)

DT	01-NOV-1981 (Rel. 20, Last sequence update)	CC	-!- stimulation of cAMP levels in the tubules.
DT	15-MAR-2004 (Rel. 43, Last annotation update)	CC	-!- SUBCELLULAR LOCATION: Secreted.
DT	Diuretic hormone (DH) (Diuretic peptide) (DP).	CC	-!- SIMILARITY: Belongs to the sauvagine/corticotropin-releasing factor/uroctasin I family.
OS	Locusta migratoria (Migratory locust).	CC	Factor/uroctasin I family.
Eukarya	Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	CC	
Neopatra	Orthopteroidea; Orthoptera; Caelifera; Acridoidea;	CC	
OC	Acridoidea; Acriidae; Oedipodinae; Locustidae; Locusta;	CC	
NCBI_TaxID	[1]	CC	
RN	SEQUENCE.	CC	
RP	TISSUE=Brain; PubMed=1654896; PMID=1654896;	CC	
RX	Lehmerg B., Ora R.B., Furuya K., King D.S., Applebaum S.W.,	CC	
RX	Kay I., Wheeler C.M., Totty N.F., Cusinato C., Patel M.,	CC	
RX	Fernicoff H.J., Schooley D.A.; "Identification of a diuretic hormone of Locusta migratoria.";	CC	
RX	Biochem. Biophys. Res. Commun. 179:1036-1041(1991).	CC	
RN	[2]	CC	
RP	SEQUENCE.	CC	
RC	TISSUE=Malpighian tubules; PubMed=163363; PMID=163363;	CC	
RX	Kay I., Wheeler C.H., Coast G.M., Totty N.F., Patel M.,	CC	
RX	Goldsborough G.J.; "Characterization of a diuretic peptide from Locusta migratoria.";	CC	
RX	Biol. Chem. Hoppe-Seyler 372:929-934(1991).	CC	
CC	-!- FUNCTION: Regulation of fluid secretion. Stimulates primary urine secretion by Malpighian tubules and causes a dose-dependent stimulation of cAMP levels in the tubules.	CC	
CC	-!- SUBCELLULAR LOCATION: Secreted.	CC	
CC	-!- SIMILARITY: Belongs to the sauvagine/corticotropin-releasing factor/uroctasin I family.	CC	
DR	PIR; JH0243; SWIQUA.	CC	
DR	InterPro; IPR000187; corticoliberin.	CC	
DR	InterPro; IPR01621; Diurtc_hormn.	CC	
DR	ProDom; PDD14750; CRF; 1.	CC	
DR	SMART; SM00039; CRF; 1.	CC	
DR	Hormone; Amidation; CRF; 1.	CC	
DR	PROSITE; PS00511; CRF; 1.	CC	
DR	PRT; Ps00511; CRF; 1.	CC	
KW	Hormone; Amidation.	CC	
FT	MOD RES 46 AA; 5364 MW; E063266EB74939763 CRC64;	CC	
SQ	SEQUENCE 46 AA; 5364 MW; E063266EB74939763 CRC64;	CC	
Qy	1 GPSL 4	CC	
Db	4 GPSL 7	CC	
RESULT 3	DIHUH_PERAM	STANDARD;	PRT; 46 AA.
ID	DIHUH_PERAM	STANDARD;	PRT; 46 AA.
AC	P41538; 32	STANDARD;	PRT; 46 AA.
DT	01-NOV-1995 (Rel. 32, Created)	STANDARD;	PRT; 46 AA.
DT	01-NOV-1995 (Rel. 32, Last sequence update)	STANDARD;	PRT; 46 AA.
DT	Diuretic hormone (DH) (Diuretic Peptide) (DP).	STANDARD;	PRT; 46 AA.
DE	Periplaneta americana (American Cockroach).	STANDARD;	PRT; 46 AA.
OS	Eukarya; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	STANDARD;	PRT; 46 AA.
OC	Neoptera; Orthopteroidea; Dictyoptera; Blattodea; Blattidae; Periplaneta.	STANDARD;	PRT; 46 AA.
OC	NCBI_TaxID=6978;	STANDARD;	PRT; 46 AA.
RN	[1]	STANDARD;	PRT; 46 AA.
RP	SEQUENCE.	STANDARD;	PRT; 46 AA.
RA	Goldsborough G.J.; "Isolation, characterization and biological activity of a CRF-related diuretic peptide from Periplaneta americana L.";	STANDARD;	PRT; 79 AA.
RT	Regul. Pept. 42:111-122(1992).	STANDARD;	PRT; 79 AA.
CC	-!- FUNCTION: Regulation of fluid secretion. Stimulates primary urine secretion by Malpighian tubules and causes a dose-dependent	STANDARD;	PRT; 79 AA.
CC	"Isolation, characterization and biological activity of a CRF-related diuretic peptide from Periplaneta americana L.";	STANDARD;	PRT; 79 AA.
DT	01-NOV-1997 (Rel. 35, Created)	STANDARD;	PRT; 79 AA.

01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-DEC-2003 (Rel. 41, Last annotation update)
 DB Methylamine utilization protein MAUD (Fragments).
 GN Methyllobacillus flagellatum.
 OS Bacteria; Proteobacteria; Betaproteobacteria; Methylophilales;
 OC Methylophilaceae; Methylobacillaceae; Methylobacillus.
 NCBI_TaxID=405;
 OX [1] _;

SEQUENCE FROM N.A.
 STRAIN=NKT / ATCC 51484 / DSM 6875 / VRM B-1610;
 RX PMID=9362636; PubMed=793847;

RA Gak E.R., Chitsoserdov A.Y., Lidstrom M.E.;
 RT "Cloning, sequencing, and mutation of a gene for azurin in
 RT Methyllobacillus flagellatum Kr.";
 RT J. Bacteriol. 177:4575-4578 (1995).

RN SEQUENCE OF 50-79 FROM N.A.
 RP STRAIN=NKT / ATCC 51484 / DSM 6875 / VRM B-1610;
 RC Zhang X., Chitsoserdov A.Y., McIntire W.S.;

RA Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases
 CC -!
 CC -! FUNCTION: MAY BE SPECIFICALLY INVOLVED IN THE PROCESSING,
 CC TRANSPORT, AND/OR MATURATION OF THE NADH BETA-SUBUNIT.
 CC -!
 CC PATHWAY: Methylamine utilization.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

CC -!
 CC -!
 CC -!

FT L37436; AAC1482.1; -;
 DR EMBL; L37427; AAC1473.1; -;
 DR EMBL; AF114265; AF03759.1; -;
 DR EMBL; AF114265; AF03759.1; -;

KW TRANSMEMBRANE.
 FT NON_CONS 50 51 MW; B4536D2B4A312CEF CRC64;
 FT POTENTIAL.
 SEQUENCE 79 AA; 8517 MW; 34536D2B4A312CEF CRC64;

Query Match Score 21; DB 1; Length 83;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPSL 4
 Db 74 GPSL 77

RESULT 7
 N19M_HUMAN
 ID N19M_HUMAN
 AC O95167;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 AC Q02371
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DB NADH:ubiquinone oxidoreductase B9 subunit (EC 1.6.5.3) (EC 1.6.99.3)
 NDPA3
 GN [1]
 OS Bos taurus (Bovine).
 OC Bivalvia; Mollusca; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae; Bovoidae;
 OC Bovidae; Bovinae; Bos.
 OX [1]
 NN SEQUENCE FROM N.A., AND SEQUENCE OF 43-73.
 RP TISSUE=Heart.
 RX MEDLINE=22389317; PubMed=1518044;
 RA Walker J.E., Arizmendi J.M., Fearnley I.M., Finel M.,
 RA Medd S.M., Filkington S.J., Runswick M.J., Skelton J.M.;
 RT "Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from
 bovine heart mitochondria. Application of a novel strategy for
 RT

Shen Y., Pan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W., Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.; "Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD4+ hematopoietic stem/progenitor cells."; *J. Genet. Res.* 10:1546-1560 (2000).

[3] TISSUE=Prostate;
RC TISSUE=Prostate;
RN MEDLINE=22388237; PubMed=1247932;
RN SEQUENCE FROM N.A.
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Bhat N.K., Altenschul S.F., Zeepberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marubina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinici P., Prange C., Raha S.S., Loquai-Labey N.A., Peters G.J., Abramson R.D., Mulahay S.J., Bosak S.S., McIwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villacon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Bouffard G.C., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grinwald J.J., Schmitz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalata U., Smailus D.E., Schnech A., Schein J.E., Jones S.J.M., Marra M.A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; *Proc. Natl. Acad. Sci. U.S.A.* 99:16899-16903 (2002).

RL FUNCTION: Transfer of electrons from NADH to the respiratory chain. The immediate electron acceptor for the enzyme is believed to be ubiquinone.

CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.

CC -!- SUBUNIT: Mammalian complex I is composed of 45 different subunits.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane.

CC -!- SIMILARITY: BELONGS TO THE COMPLEX I B9 SUBUNIT FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; AF044955; AAD05420.1; -.
CC DR EMBL; AF070653; AAD05959.1; -.
CC DR PIR; JB0379; JB1379.
CC DR Genew; HGNC:7688; NDUF4A3.
CC DR MIM; 603832; -.
CC DR GO:001377; F:NADH dehydrogenase (ubiquinone) activity; TAS.
KW Oxidoreductase; Ubiquinone; NAD; Mitochondrion; Transmembrane.
FT TRANSMEM 19 39 POTENTIAL.
SEQUENCE 84 AA; 9279 MW; 38B2A96D7A05D31 CRC64;

Best Local Similarity 100.0%; Score 21; DB 1; Length 84;
Matches 4; Conservative 100 %; Pred. No. 2.3e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPSL 4
Db 75 GPSL 78

RESULT 8
N19M_MOUSE STANDARD; PRT; 84 AA.
ID N19M_MOUSE STANDARD; PRT; 84 AA.
AC P00899;
ID Q9CQ91; 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)

28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NADH-ubiquinone oxidoreductase B9 subunit (EC 1.6.5.3) (EC 1.6.99.3)
DB (Complex I-B9) (CI-B9).
GN NDUF4A3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TAXID=10090;
RN [1];
RN RN RP
RN SEQUENCE FROM N.A.
RC STAIN=57BL/6U; TISSUE=Heart, and Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawa J., Shinagawa A., Shibara K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiysawa H., Kondo S., Yamamoto I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Gustincich S., Hill D., Fletcher C., Hefmann M., Kamiya M., Lee N.H., Raasch W., Gaasterland T., Gissi C., King B., Kochiwa H., Kiehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Orido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Borjunga N., Carninci P., de Bonaldo M.P., RA Brownstein M.J., Bult C., Fletcher C., Gojobori M., Garibaldi M., RA Gustincich S., Hill D., Hame D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarilli J., Nombara S., RA Noda M., Ring B., Rodriguez I., Sakamoto N., RA Sasaki H., Sato K., Schoenbach C., Suya T., Shiba T., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wining L., RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsu S., RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
CC -!- FUNCTION: Transfer of electrons from NADH to the respiratory chain. The immediate electron acceptor for the enzyme is believed to be ubiquinone (By similarity).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- SUBUNIT: Mammalian complex I is composed of 45 different subunits.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I B9 SUBUNIT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC DR EMBL; AK019059; BAB31545.1; -.
CC DR EMBL; AK03133; BAB22593.1; -.
CC DR MGI; MGJ:1911341; Ndufa3.
KW Oxidoreductase; Ubiquinone; NAD; Mitochondrion; Transmembrane.
FT TRANSMEM 19 39 POTENTIAL.
SQ SEQUENCE 84 AA; 9331 MW; B05E598217D95A92 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.3e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPSL 4
Db 75 GPSL 78

RESULT 9
CY22_RHOPT STANDARD; PRT; 96 AA.
ID CY22_RHOPT STANDARD; PRT; 96 AA.
AC P00899;
ID Q9CQ91; 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT	28-FEB-2003 (Ref. 41, Last annotation update)	DE	Cytochrome C2, 160-2.
DE	30S ribosomal protein S6.	OS	Rhodospirillum molischianum.
RSF	OR ML2685 OR MdB1913-21C.	CC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
GS	Mycobacterium leprae	OC	Rhodospirillaceae; Phaeospirillium.
OC	Bacteria; Actinobacteridae; Actinomycetales;	NCBI_TaxID=1083 ;	
OC	Corynebacterineae; Mycobacterium.	[1]	
OX		RP	SEQUENCE.
NCBI_TaxID=1769;		RA	Ambler R.P.; Meyer T.E.; Bartsch R.G.; Kamen M.D.;
RN	SEQUENCE FROM N.A. PubMed=969512;	RL	Unpublished results, cited by:
RX	Fsihi H., de Rossi E., Salazar L., Cantoni R., Labo M., Riccardi G., Takifi H.E., Eiglmeier K., Bergh S., Cole S.T.; "Gene arrangement and organization in an approximately 76 kb fragment encompassing the oric region of the chromosome of Mycobacterium leprae"; Microbiology 142:3147-3161.(1996).	RL	RL (In) Matsubara H., Yamakawa T. (eds.); Evolution of protein molecules, pp 311-322, Japan Scientific Societies Press/Center for Academic Publications, Tokyo (1978).
RN	SEQUENCE FROM N.A. MEDLINE=97124199 ; PubMed=969512;	RA	[2]
RX	RNA STRAIN-TN; MEDLINE=21122732; PubMed=11234002;	RA	SEQUENCE.
RA	RA	RA	Ambler R.P.;
RA	RA	RA	Submitted (JUN-1977) to the PIR data bank.
RA	RA	RA	CC - I - FUNCTION: CYTOCHROME C2 IS FOUND MAINLY IN PURPLE, NONSULFUR, PHOTOSYNTHETIC BACTERIA WHERE IT FUNCTIONS AS THE ELECTRON DONOR TO THE OXIDIZED BACTERIOCHLOROPHYLL IN THE PHOTOPHOSPHORYLATION PATHWAY. HOWEVER, IT MAY ALSO HAVE A ROLE IN THE RESPIRATORY CHAIN AND IS FOUND IN SOME NONPHOTOSYNTHETIC BACTERIA.
RA	RA	RA	DR A00080; CCQFM2.
RA	RA	RA	DR P00004; IWEJ.
RA	RA	RA	DR IP003088; Cyt CI.
RA	RA	RA	DR InterPro; IP002327; Cyt-CIAB.
RA	RA	RA	DR InterPro; IP000345; Cyt-C_heme_BS.
RA	RA	RA	DR PF00034; cytochrome c; 1.
RA	RA	RA	DR PR00034; CYTCHEMCIAB.
RA	RA	RA	DR Prodrom; PD000375; Cyt CIAB; 1.
RA	RA	RA	DR PROSITE; PS00190; CYTOCHROME C; 1.
RA	RA	KW	KW Electron transport; Photosynthesis; Heme.
RA	RA	FT	FT BINDING 10 HEME (COVALENT).
RA	RA	FT	FT BINDING 13 HEME (COVALENT).
RA	RA	FT	FT METAL 14 IRON (HEME AXIAL LIGAND).
RA	RA	FT	FT METAL 75 IRON (HEME AXIAL LIGAND).
RA	RA	SQ	Query Match Score 21; DB 1; Length 97; Best Local Similarity 100.0%; Pred. No. 2.7e+02; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RA	RA	Qy	Qy 1 GPSL 4
RA	RA	Db	Db 24 GPSL 27
CC	CC	CC	RESULT 13
CC	CC	CC	CY21_RHOFU STANDARD; PRT; 99 AA.
CC	CC	CC	ID -CY21_RHOFU AC P00086; (Rel. 01, Created)
CC	CC	CC	DT 21-JUL-1986 (Rel. 01, Last sequence update)
CC	CC	CC	DT 21-JUL-1986 (Rel. 01, Last annotation update)
CC	CC	CC	DT 15-DEC-1998 (Rel. 37, Last annotation update)
CC	CC	CC	DB Cytochrome C2, 160-1.
CC	CC	CC	OS Rhodospirillum fulvum.
CC	CC	CC	OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
CC	CC	CC	NCBI_TaxID=1082; [1]
CC	CC	CC	RP SEQUENCE.
CC	CC	CC	RA Ambler R.P.; Meyer T.E.; Bartsch R.G.; Kamen M.D.;
CC	CC	CC	RL Unpublished results, cited by:
CC	CC	CC	RL (In) Matsubara H., Yamakawa T. (eds.); Evolution of protein molecules, pp 311-322, Japan Scientific Societies Press/Center for Academic Publications, Tokyo (1978).
CC	CC	CC	[2]
CC	CC	CC	RP
CC	CC	CC	RA
CC	CC	CC	RL Submitted (JUN-1977) to the PIR data bank.
CC	CC	CC	CC - I - FUNCTION: CYTOCHROME C2 IS FOUND MAINLY IN PURPLE, NONSULFUR, PHOTOSYNTHETIC BACTERIA WHERE IT FUNCTIONS AS THE ELECTRON DONOR
CC	CC	CC	CC

CC TO THE OXIDIZED BACTERIOCHLOROPHYLL IN THE PHOTOPHOSPHORYLATION
CC PATHWAY. HOWEVER, IT MAY ALSO HAVE A ROLE IN THE RESPIRATORY CHAIN
CC AND IS FOUND IN SOME NONPHOTOSYNTHETIC BACTERIA.

DR PIR: A00078; CCQFNF.

DR HSSP: P00045; 1YES.

DR InterPro: IPRO03058; Cyt-CI.

DR InterPro: IPRO02327; Cyt-CIAB.

DR InterPro: IPRO00345; CytC_heme_BS.

DR Pfam: PF00341; cytochrome_c_1.

DR PRINTS: PR00044; CYTOCHRMBCIAB.

DR PROSITE: PS000375; Cyt_CIAB_1.

DR PROSITE: PS00190; CYTOCHROMB_C_1.

DR KW Electron transport; Photosynthesis; Heme.

DR FT BINDING 10 HEME (COVALENT).

DR FT BINDING 13 HEME (COVALENT).

DR FT METAL 14 IRON (HEME AXIAL LIGAND).

DR FT METAL 75 IRON (HEME AXIAL LIGAND).

DR SQ SEQUENCE 99 AA; 10288 MW; DECAL125B7457B022 CRC64;

DR

CC TO THE OXIDIZED BACTERIOCHLOROPHYLL IN THE PHOTOPHOSPHORYLATION
CC PATHWAY. HOWEVER, IT MAY ALSO HAVE A ROLE IN THE RESPIRATORY CHAIN
CC AND IS FOUND IN SOME NONPHOTOSYNTHETIC BACTERIA.

DR PIR: A00078; CCQFNF.

DR HSSP: P00045; 1YES.

DR InterPro: IPRO03058; Cyt-CI.

DR InterPro: IPRO02327; Cyt-CIAB.

DR InterPro: IPRO00345; CytC_heme_BS.

DR Pfam: PF00341; cytochrome_c_1.

DR PRINTS: PR00044; CYTOCHRMBCIAB.

DR PROSITE: PS000375; Cyt_CIAB_1.

DR PROSITE: PS00190; CYTOCHROME_C_1.

DR ELECTRON TRANSPORT; PHOTOSYNTHESIS; Heme (COVALENT).

DR FT BINDING 11 HEME (COVALENT).

DR FT BINDING 14 HEME (COVALENT).

DR FT METAL 15 IRON (HEME AXIAL LIGAND).

DR FT METAL 76 IRON (HEME AXIAL LIGAND).

DR SQ SEQUENCE 100 AA; 10248 MW; 6AF5FB2E3BF19D0 CRC64;

DR

CC TO THE OXIDIZED BACTERIOCHLOROPHYLL IN THE PHOTOPHOSPHORYLATION
CC PATHWAY. HOWEVER, IT MAY ALSO HAVE A ROLE IN THE RESPIRATORY CHAIN
CC AND IS FOUND IN SOME NONPHOTOSYNTHETIC BACTERIA.

DR PIR: A00078; CCQFNF.

DR HSSP: P00045; 1YES.

DR InterPro: IPRO03058; Cyt-CI.

DR InterPro: IPRO02327; Cyt-CIAB.

DR InterPro: IPRO00345; CytC_heme_BS.

DR Pfam: PF00341; cytochrome_c_1.

DR PRINTS: PR00044; CYTOCHRMBCIAB.

DR PROSITE: PS000375; Cyt_CIAB_1.

DR PROSITE: PS00190; CYTOCHROME_C_1.

DR ELECTRON TRANSPORT; PHOTOSYNTHESIS; Heme (COVALENT).

DR FT BINDING 11 HEME (COVALENT).

DR FT BINDING 14 HEME (COVALENT).

DR FT METAL 15 IRON (HEME AXIAL LIGAND).

DR FT METAL 76 IRON (HEME AXIAL LIGAND).

DR SQ SEQUENCE 100 AA; 10248 MW; 6AF5FB2E3BF19D0 CRC64;

DR

CC TO THE OXIDIZED BACTERIOCHLOROPHYLL IN THE PHOTOPHOSPHORYLATION
CC PATHWAY. HOWEVER, IT MAY ALSO HAVE A ROLE IN THE RESPIRATORY CHAIN
CC AND IS FOUND IN SOME NONPHOTOSYNTHETIC BACTERIA.

DR PIR: A00078; CCQFNF.

DR HSSP: P00045; 1YES.

DR InterPro: IPRO03058; Cyt-CI.

DR InterPro: IPRO02327; Cyt-CIAB.

DR InterPro: IPRO00345; CytC_heme_BS.

DR Pfam: PF00341; cytochrome_c_1.

DR PRINTS: PR00044; CYTOCHRMBCIAB.

DR PROSITE: PS000375; Cyt_CIAB_1.

DR PROSITE: PS00190; CYTOCHROME_C_1.

DR ELECTRON TRANSPORT; PHOTOSYNTHESIS; Heme (COVALENT).

DR FT BINDING 11 HEME (COVALENT).

DR FT BINDING 14 HEME (COVALENT).

DR FT METAL 15 IRON (HEME AXIAL LIGAND).

DR FT METAL 76 IRON (HEME AXIAL LIGAND).

DR SQ SEQUENCE 100 AA; 10248 MW; 6AF5FB2E3BF19D0 CRC64;

DR

CC TO THE OXIDIZED BACTERIOCHLOROPHYLL IN THE PHOTOPHOSPHORYLATION
CC PATHWAY. HOWEVER, IT MAY ALSO HAVE A ROLE IN THE RESPIRATORY CHAIN
CC AND IS FOUND IN SOME NONPHOTOSYNTHETIC BACTERIA.

DR PIR: A00078; CCQFNF.

DR HSSP: P00045; 1YES.

DR InterPro: IPRO03058; Cyt-CI.

DR InterPro: IPRO02327; Cyt-CIAB.

DR InterPro: IPRO00345; CytC_heme_BS.

DR Pfam: PF00341; cytochrome_c_1.

DR PRINTS: PR00044; CYTOCHRMBCIAB.

DR PROSITE: PS000375; Cyt_CIAB_1.

DR PROSITE: PS00190; CYTOCHROME_C_1.

DR ELECTRON TRANSPORT; PHOTOSYNTHESIS; Heme (COVALENT).

DR FT BINDING 11 HEME (COVALENT).

DR FT BINDING 14 HEME (COVALENT).

DR FT METAL 15 IRON (HEME AXIAL LIGAND).

DR FT METAL 76 IRON (HEME AXIAL LIGAND).

DR SQ SEQUENCE 100 AA; 10248 MW; 6AF5FB2E3BF19D0 CRC64;

DR

Db 4 GPSL 7

RESULT 3

PN0584

N1: Alternative names: oxygen oxidoreductase; tetrahydropteridine tyrosine hydroxylase
 C1: Species: Gorilla gorilla (gorilla)

C2: Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000

C3: Accession: PN0584

R1: Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.
 Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A1: Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A2: Reference number: PN0575; MUID:93371398; PMID:7689834

A3: Accession: PN0584

A4: Molecule type: Genomic RNA

A5: Residues: 1-27 <ICH>

A6: Cross-references: GB:Li14796

A7: Experimental source: lymphocytes of peripheral blood

C1: Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine biosynthesis

C2: Superfamily: Phenylalanine 4-monoxygenase

C3: Keywords: biotin; monooxygenase; oxidoreductase

Qy 1 GPSL 4

Db 4 GPSL 7

RESULT 6

S20771

Ig heavy chain V region (VH4, IN1P24) - human (fragment)

C1: Species: Homo sapiens (man)

C2: Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 23-Jul-1999

C3: Accession: S20771

R1: Mortari, F.; Wang, J.; Schroeder, H. W.

submitted to the EMBL Data Library, April 1992

A1: Description: Analysis of human cord blood Ig heavy chain IgA and IgG repertoire.

A2: Reference number: S20764

A3: Accession: S20771

A4: Molecule type: DNA

A5: Residues: 1-32 <MoR>

A6: Cross-references: EMBL:Z11955; NID:933871; PIDN:CAA78012_1; PMID:933872

C1: Superfamily: immunoglobulin V region; immunoglobulin homology

C2: Keywords: heterotetramer; immunoglobulin

RESULT 4

PN0585

Tyrosine 3-monoxygenase (EC 1.14.16.2) type 3 and 4 - orangutan (fragment)

C1: Species: Pongo pygmaeus (orangutan)

C2: Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000

C3: Accession: PN0585

R1: Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.
 Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A1: Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A2: Reference number: PN0575; MUID:93371398; PMID:7689834

A3: Accession: PN0585

A4: Molecule type: Genomic RNA

A5: Residues: 1-27 <ICH>

A6: Cross-references: GB:Li14799

A7: Experimental source: lymphocytes of peripheral blood

C1: Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine biosynthesis

C2: Superfamily: Phenylalanine 4-monoxygenase

C3: Keywords: biotin; monooxygenase; oxidoreductase

Qy 1 GPSL 4

Db 16 GPSL 19

RESULT 7

C25159

Hypothetical protein 3 4.2K sin - Bacillus subtilis

C1: Species: Bacillus subtilis

C2: Date: 16-Aug-1998 #sequence_revision 16-Aug-1998 #text_change 16-Dec-1998

C3: Accession: C25159

R1: Gaur, N.K.; Dubnau, E.; Smith, I.

J. Bacteriol. 168, 860-869, 1986

A1: Title: Characterization of a cloned *Bacillus subtilis* gene that inhibits sporulation

A2: Reference number: A91827; MUID:87056981; PMID:3096962

A3: Accession: C25159

A4: Molecule type: DNA

A5: Residues: 1-35 <GAU>

Qy 1 GPSL 4

Db 16 GPSL 19

Query Match Score 21; DB 2; Length 32; Best Local Similarity 100.0%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;

Query Match Score 21; DB 2; Length 32; Best Local Similarity 100.0%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;

Query Match Score 21; DB 2; Length 35; Best Local Similarity 100.0%; Pred. No. 2.1e+02; Mismatches 0; Indels 0; Gaps 0;

Query Match Score 21; DB 2; Length 35; Best Local Similarity 100.0%; Pred. No. 2.1e+02; Mismatches 0; Indels 0; Gaps 0;

RESULT 8

SWLQDA

African migratory locust

C1: Species: Locusta migratoria migratorioides (African migratory locust)

C2: Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 03-Nov-2003

C3: Accession: JH0243; S17864; A23702

Qy 1 GPSL 4

Db 4 GPSL 7

RESULT 5

PN0586

Tyrosine 3-monoxygenase (EC 1.14.16.2) type 3 and 4 - common gibbon (fragment)

C1: Species: Hylobates lar (common gibbon, white-handed gibbon)

C2: Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000

C3: Accession: PN0586

R1: Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.
 Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A1: Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A2: Reference number: PN0575; MUID:93371398; PMID:7689834

B;Lehmlberg, E.; Ota, R.B.; Furuya, K.; King, D.S.; Applebaum, S.W.; Ferenz, H.J.; School Biochem. Biophys. Res. Commun. 179, 1016-1011, 1991
A;Title: Identification of a diuretic hormone of *Locusta migratoria*.
A;Reference number: JH0243; MUID:91378568; PMID:1654856
A;Accession: JH0243
A;Molecule type: protein
A;Residues: 1-46 <R22>
A;Experimental source: brain and corpora cardiaca
R;Kay, I.; Wheeler, C.H.; Coast, G.M.; Totty, N.P.; Cusinato, O.; Patel, M.; Goldsworthy, P.; Hoppe-Seyler, 372, 929-934, 1991
A;Title: Characterization of a diuretic Peptide from *Locusta migratoria*.
A;Reference number: S17864; MUID:92126231; PMID:1663363
A;Molecule type: protein
A;Residues: 1-46 <R2Y>
A;Note: Species designated as *Locusta migratoria*
C;Comment: This hormone stimulates urine production by Malpighian tubules and elevates blood pressure
C;Superfamily: diuretic hormone; diuretic hormone homology
C;Keywords: amidated carboxyl end; diuretic hormone homology <DEH>
F;2-46/Domain: diuretic hormone homology
F;46/Modified site: amidated carboxyl end (Ile) #status experimental
Query Match Score 21; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPSL 4
Db 4 GPSL 7

RESULT 9
A48542
CRF-related diuretic peptide - American cockroach
C;Species: *Periplaneta americana* (American cockroach)
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 03-Nov-2003
C;Accession: A48542
R;Kay, I.; Patel, M.; Coast, G.M.; Totty, N.F.; Mallet, A.I.; Goldsworthy, G.J.
Regui, P.; PepB, 42, 111-122, 1992
A;Title: Isolation, characterization and biological activity of a CRP-related diuretic
A;Reference number: A48542; MUID:31174045; PMID:1337734
A;Accession: A48542
A;Status: Preliminary
A;Molecule type: protein
A;Residues: 1-46 <R2Y>
A;Note: sequence extracted from NCBI backbone (NCBIP:125548)
C;Superfamily: diuretic hormone; diuretic hormone homology
F;2-46/Domain: diuretic hormone homology <DEH>
Query Match Score 21; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPSL 4
Db 4 GPSL 7

RESULT 10
F90584
Hypothetical protein MYFU 5820 [imported] - *Mycoplasma pulmonis* (strain UAB CTIP)
C;Species: *Mycoplasma pulmonis*
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C;Accession: F90584
R;Chambaud, I.; Heilig, R.; Perris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.; Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulm*
A;Reference number: A99512; MUID:21226715; PMID:11353084
A;Accession: F90584
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-50 <KUR>
A;Cross-references: GB:AL445566; PID:gi14089997; PID:CAIC13755.1; GSPDB:GN00153

A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYFU 5820
A;Genetic code: SGCG3
Query Match Score 21; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPSL 4
Db 41 GPSL 44

RESULT 11
C38230
inorganic diphosphatase (EC 3.6.1.1) 1 - beet (fragments)
C;Species: Beta vulgaris (beet)
C;Accession: C38230 #sequence_revision 20-Aug-1999 #text_change 27-Oct-2003
C;Accession: C38230; B38230; F38230
R;Sarafian, V.; Kim, Y.; Poole, R.J.; Rea, P.A.; Proc. Natl. Acad. Sci. U.S.A., 89, 1775-1779, 1992
A;Title: Molecular cloning and sequence of cDNA encoding the pyrophosphate-energized v A;Reference number: A38230; MUID:92179265; PMID:1311852
A;Accession: C38230
A;Molecule type: protein
A;Residues: 1-16 <SR1>
A;Note: sequence extracted from NCBI backbone (NCBIP:871922)
A;Accession: B38230
A;Molecule type: protein
A;Residues: 1-28 <SR2>
A;Note: sequence extracted from NCBI backbone (NCBIP:871941)
A;Accession: F38230
A;Molecule type: protein
A;Residues: 29-32 <SR3>
A;Experimental source: cultivar Detroit Dark
A;Note: sequence extracted from NCBI backbone (NCBIP:87195)
C;Superfamily: H(+)-translocating inorganic pyrophosphatase
C;Keywords: hydrolase
Query Match Score 21; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPSL 4
Db 45 GPSL 48

RESULT 12
S35628
myosin heavy chain - rabbit (fragment)
C;Species: *Oryctolagus cuniculus* (domestic rabbit)
C;Accession: S35628
R;Rabij, P.; Nucleic Acids Res. 21, 1467-1471, 1993
A;Reference number: S35628; MUID:93219127; PMID:8464739
A;Accession: S35628
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-54 <BAB>
A;Cross references: BMBL:X70964; NID:g297108; PID:9939958
C;Superfamily: myosin heavy chain; myosin motor domain homology
F;1-54/Domain: myosin motor domain homology (fragment) <MMOT>
Query Match Score 21; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 GPSL 26

RESULT 13
A;Residues: 1-56 <ANS>A;Cross-references: EMBL:X73008; NID:9295875; PIDN:CAA51495.1; PID:9939938
R;Kettleborough, C.A.; Saldana, J.; Ansell, K.H.; Bendig, M.M.C;Date: 11-Jan-2000 #sequence_change 11-Jan-2000
C;Accession: T31181A;Residue number: 206-211; 1993
A;Submitted to: the EMBL Data Library, July 1998A;Description: Complete sequence of a 184 kb catabolic plasmid from *Sphingomonas aromatic*A;Reference number: Z20932
A;Accession: T31181A;Status: Preliminary; translated from GB/EMBL/DBJ
A;Residues: 1-54 <RDN>A;Cross-references: EMBL:AF079317; NID:93378261; PIDN:AA03905.1
C;Genetics:A;Genome: Plasmid pNL1
A;Note: orf505

Query Match 100.0% Score 21; DB 2; Length 56;

Best Local Similarity 100.0% Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: March 10, 2004, 09:16:43

Job time : 3.30395 secs

Qy 1 GPSL 4

Db 12 GPSL 15

RESULT 14

T03658
Phosphoenolpyruvate carboxylase (EC 4.1.1.31) - maize (fragment).

C;Species: Zea mays (maize)

C;Date: 24-Mar-1999 #sequence_change 24-Mar-1999

C;Accession: T03658

P;Yangisawa, S.; Iwui, K.; Yamaguchi, Y.; Shigeeda, K.; Katsumi, H.
PEBS Dect. 229, 107-110, 1988

A;Title: Further analysis of cDNA clones for maize phosphoenolpyruvate carboxylase involved in RNA at multiple sites in vivo.

A;Reference number: S00348; NID:D:88152202; PMID:2894322

A;Accession: T03658

A;Status: Preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-56 <YAN>

A;Cross-references: EMBL:X07171; NID:922414; PIDN:CAA30155.1; PID:9929919
C;Function:

A;Description: catalyzes the irreversible beta-carboxylation of phosphoenolpyruvate by b

A;Pathway: carbon dioxide fixation

C;Superfamily: Phosphoenolpyruvate carboxylase

C;Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoprotein

Query Match 100.0% Score 21; DB 2; Length 56;

Best Local Similarity 100.0% Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPSL 4

Db 28 GPSL 36

RESULT 15

S33380
19 heavy chain V region - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 19-Mar-1997 #sequence_change 29-Aug-1997

C;Accession: S36380; S33395

R;Ansell, K.H.

A;Submitted to the EMBL Data Library, April 1993

A;Reference number: S36376

A;Accession: S36380

A;Molecule type: mRNA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 15.7198 Seconds

(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-3

Perfect score: 95

Sequence: 1 DEKKIARMEKASSVFNVNS 20

Scoring table: BLOSUM62

Gapext 0.5

Searched: 809712 seqs, 211153259 residues

Total number of hits satisfying chosen parameters:

809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database :

```

1: /cgns_6/prodata/2/pubpaas/US07_PUBCOMB.pep:*
2: /cgns_6/prodata/2/pubpaas/PCT_PUBCOMB.pep:*
3: /cgns_6/prodata/2/pubpaas/US08_PUBCOMB.pep:*
4: /cgns_6/prodata/2/pubpaas/US09_PUBCOMB.pep:*
5: /cgns_6/prodata/2/pubpaas/US10_PUBCOMB.pep:*
6: /cgns_6/prodata/2/pubpaas/PCRTS_PUBCOMB.pep:*
7: /cgns_6/prodata/2/pubpaas/US08_NN_PUBCOMB.pep:*
8: /cgns_6/prodata/2/pubpaas/US09_PUBCOMB.pep:*
9: /cgns_6/prodata/2/pubpaas/US09B_PUBCOMB.pep:*
10: /cgns_6/prodata/2/pubpaas/US09C_PUBCOMB.pep:*
11: /cgns_6/prodata/2/pubpaas/US09_T_PUBCOMB.pep:*
12: /cgns_6/prodata/2/pubpaas/US10_N_PUBCOMB.pep:*
13: /cgns_6/prodata/2/pubpaas/US10A_PUBCOMB.pep:*
14: /cgns_6/prodata/2/pubpaas/US10B_PUBCOMB.pep:*
15: /cgns_6/prodata/2/pubpaas/US10C_PUBCOMB.pep:*
16: /cgns_6/prodata/2/pubpaas/US60_N_PUBCOMB.pep:*
17: /cgns_6/prodata/2/pubpaas/US60_T_PUBCOMB.pep:*
18: /cgns_6/prodata/2/pubpaas/US60_U_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	20	9 US-09-848-834A-3	Sequence 3, Appl1
2	95	100.0	33	9 US-09-848-834A-12	Sequence 12, Appl1
3	95	100.0	36	9 US-09-848-834A-16	Sequence 16, Appl1
4	95	100.0	51	9 US-09-848-834A-20	Sequence 20, Appl1
5	89	93.7	19	14 US-10-239.313A-54	Sequence 54, Appl1
6	89	93.7	21	10 US-09-932.165.1482	Sequence 1482, Appl1
7	89	93.7	21	10 US-09-935.384-10	Sequence 710, Appl1
8	89	93.7	21	10 US-09-942.052-711	Sequence 711, Appl1
9	89	93.7	21	14 US-10-001.469.1404	Sequence 1404, Appl1
10	89	93.7	21	14 US-10-128.711.97	Sequence 97, Appl1
11	89	93.7	21	14 US-10-116.118.33	Sequence 33, Appl1
12	89	93.7	21	14 US-10-062.109.761	Sequence 761, Appl1
13	89	93.7	21	14 US-10-005.480.761	Sequence 761, Appl1
14	89	93.7	21	14 US-10-277.292.652	Sequence 652, Appl1
15	89	93.7	21	15 US-10-291.241-26	Sequence 26, Appl1

Sequence 652, Appl1
Sequence 2583, Appl1
Sequence 25, Appl1
Sequence 2, Appl1
Sequence 52, Appl1
Sequence 2, Appl1
Sequence 44, Appl1
Sequence 422, Appl1
Sequence 39, Appl1
Sequence 23, Appl1
Sequence 15, Appl1
Sequence 17, Appl1
Sequence 39, Appl1
Sequence 31, Appl1
Sequence 96, Appl1
Sequence 53, Appl1
Sequence 3, Appl1
Sequence 36, Appl1
Sequence 38, Appl1
Sequence 39, Appl1
Sequence 41, Appl1
Sequence 60, Appl1
Sequence 64, Appl1
Sequence 67, Appl1
Sequence 59, Appl1
Sequence 65, Appl1
Sequence 66, Appl1
Sequence 14, Appl1
Sequence 26, Appl1

ALIGNMENTS

RESULT 1
US-09-848-834A-3
; Sequence 3, Application US/09848834A
; Patent No. US20076416A1
; GENERAL INFORMATION:
; APPLICANT: Abtov Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102885-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1) (20)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Plamodium
; US-09-848-834A-3

Query Match 100.0%; Score 95; DB 9; Length: 20;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEKKIARMEKASSVFNVNS 20
Db 1 DEKKIARMEKASSVFNVNS 20

RESULT 2
US-09-848-834A-12
; Sequence 12, Application US/09848834A
; Patent No. US20076416A1

GENERAL INFORMATION:
 APPLICANT: Aphton Corporation
 FILE REFERENCE: 110265-0047
 CURRENT APPLICATION NUMBER: US/09/848, 834A
 CURRENT FILING DATE: 2001-05-04
 PRIOR APPLICATION NUMBER: 60/202, 328
 PRIOR FILING DATE: 2000-05-05
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 12
 LENGTH: 33
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 378-398 of the Plasmodium falciparum CSP protein linked by a spacer to amino acid sequence 2-10 of the GnRH hormone

NAME/KEY: PEPTIDE
 LOCATION: (11).-(16)
 OTHER INFORMATION: Spacer peptide

NAME/KEY: PEPTIDE
 LOCATION: (17).-(36)
 OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria protein falciparum circumsporozoite protein

OTHER INFORMATION: (CSP) Protein
 US-09-848-834A-16

Query Match Score 100.0%; DB 9; Length 36;
 Best Local Similarity 100.0%; Pred. No. 5.Se-08;
 Matches 20; Conservative 0; Mismatches 0;
 Indels 0; Gaps 0;

RESULT 4
 US-09-848-834A-20
 Sequence 20, Application US/09848834A
 Patent No. US20020076416A1
 GENERAL INFORMATION:
 APPLICANT: Aphton Corporation
 TITLE OF INVENTION: Chimeric Peptide Immunogens
 FILE REFERENCE: 1102865-0047
 CURRENT APPLICATION NUMBER: US/09/848, 834A
 PRIORITY FILING DATE: 2001-05-04
 PRIOR APPLICATION NUMBER: 60/202, 328
 PRIOR FILING DATE: 2000-05-05
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 20
 LENGTH: 51
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human GnRH linked by a spacer to amino acid sequence 378-398 of Plasmodium falciparum circumsporozoite (CSP) protein

NAME/KEY: MOD-RES
 LOCATION: (1).-(1)
 OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: MOD-RES
 LOCATION: (51).-(51)
 OTHER INFORMATION: Amidated glycine or glycaminide

NAME/KEY: PEPTIDE
 LOCATION: (11).-(10)
 OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone

NAME/KEY: PEPTIDE
 LOCATION: (11).-(16)
 OTHER INFORMATION: Spacer peptide

NAME/KEY: PEPTIDE
 LOCATION: (17).-(36)
 OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum circumsporozoite (CSP) protein

NAME/KEY: PEPTIDE
 LOCATION: (37).-(42)
 OTHER INFORMATION: Spacer peptide

NAME/KEY: PEPTIDE
 LOCATION: (43).-(51)
 OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone

US-09-848-834A-15
 Sequence 16, Application US/09848834A
 Patent No. US20020076416A1
 GENERAL INFORMATION:
 APPLICANT: Aphton Corporation
 TITLE OF INVENTION: Chimeric Peptide Immunogens
 FILE REFERENCE: 1102865-0047
 CURRENT APPLICATION NUMBER: US/09/848, 834A
 CURRENT FILING DATE: 2001-05-04
 PRIOR APPLICATION NUMBER: 60/202, 328
 PRIOR FILING DATE: 2000-05-05
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 16
 LENGTH: 36
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the RH hormone linked by a spacer to amino acid sequence 378-398 of the Plasmodium falciparum circumsporozoite (CSP) protein

NAME/KEY: MOD-RES
 LOCATION: (1).-(1)
 OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: PEPTIDE
 LOCATION: (1).-(10)

Query Match Score 100.0%; DB 9; Length 51;
 Best Local Similarity 100.0%; Pred. No. 8e-08;
 Matches 20; Conservative 0; Mismatches 0;
 Indels 0; Gaps 0;

RESULT 3
 US-09-848-834A-20
 Sequence 20, Application US/09848834A
 Patent No. US20020076416A1
 GENERAL INFORMATION:
 APPLICANT: Aphton Corporation
 TITLE OF INVENTION: Chimeric Peptide Immunogens
 FILE REFERENCE: 1102865-0047
 CURRENT APPLICATION NUMBER: US/09/848, 834A
 CURRENT FILING DATE: 2001-05-04
 PRIOR APPLICATION NUMBER: 60/202, 328
 PRIOR FILING DATE: 2000-05-05
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 16
 LENGTH: 36
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 378-398 of the Plasmodium falciparum circumsporozoite (CSP) protein

NAME/KEY: MOD-RES
 LOCATION: (1).-(1)
 OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: PEPTIDE
 LOCATION: (1).-(10)

Query Match Score 100.0%; DB 9; Length 51;
 Best Local Similarity 100.0%; Pred. No. 8e-08;
 Matches 20; Conservative 0; Mismatches 0;
 Indels 0; Gaps 0;

RESULT 2
 US-09-848-834A-20
 Sequence 20, Application US/09848834A
 Patent No. US20020076416A1
 GENERAL INFORMATION:
 APPLICANT: Aphton Corporation
 TITLE OF INVENTION: Chimeric Peptide Immunogens
 FILE REFERENCE: 1102865-0047
 CURRENT APPLICATION NUMBER: US/09/848, 834A
 CURRENT FILING DATE: 2001-05-04
 PRIOR APPLICATION NUMBER: 60/202, 328
 PRIOR FILING DATE: 2000-05-05
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 16
 LENGTH: 36
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Plasmodium falciparum circumsporozoite (CSP) protein

NAME/KEY: MOD-RES
 LOCATION: (1).-(1)
 OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: PEPTIDE
 LOCATION: (1).-(10)

Query Match Score 100.0%; DB 9; Length 51;
 Best Local Similarity 100.0%; Pred. No. 8e-08;
 Matches 20; Conservative 0; Mismatches 0;
 Indels 0; Gaps 0;

RESULT 1
 US-09-848-834A-20
 Sequence 20, Application US/09848834A
 Patent No. US20020076416A1
 GENERAL INFORMATION:
 APPLICANT: Aphton Corporation
 TITLE OF INVENTION: Chimeric Peptide Immunogens
 FILE REFERENCE: 1102865-0047
 CURRENT APPLICATION NUMBER: US/09/848, 834A
 CURRENT FILING DATE: 2001-05-04
 PRIOR APPLICATION NUMBER: 60/202, 328
 PRIOR FILING DATE: 2000-05-05
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 16
 LENGTH: 36
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Plasmodium falciparum circumsporozoite (CSP) protein

NAME/KEY: MOD-RES
 LOCATION: (1).-(1)
 OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: PEPTIDE
 LOCATION: (1).-(10)

Query Match Score 100.0%; DB 9; Length 51;
 Best Local Similarity 100.0%; Pred. No. 8e-08;
 Matches 20; Conservative 0; Mismatches 0;
 Indels 0; Gaps 0;

RESULT 0
 US-09-848-834A-20
 Sequence 20, Application US/09848834A
 Patent No. US20020076416A1
 GENERAL INFORMATION:
 APPLICANT: Aphton Corporation
 TITLE OF INVENTION: Chimeric Peptide Immunogens
 FILE REFERENCE: 1102865-0047
 CURRENT APPLICATION NUMBER: US/09/848, 834A
 CURRENT FILING DATE: 2001-05-04
 PRIOR APPLICATION NUMBER: 60/202, 328
 PRIOR FILING DATE: 2000-05-05
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO: 16
 LENGTH: 36
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Plasmodium falciparum circumsporozoite (CSP) protein

NAME/KEY: MOD-RES
 LOCATION: (1).-(1)
 OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: PEPTIDE
 LOCATION: (1).-(10)

Query Match Score 100.0%; DB 9; Length 51;
 Best Local Similarity 100.0%; Pred. No. 8e-08;
 Matches 20; Conservative 0; Mismatches 0;
 Indels 0; Gaps 0;

RESULT 5
 US-10-339-313A-54
 Sequence 54, Application US/102393913A
 GENERAL INFORMATION No. US2003017528541
 APPLICANT: KLINGUBER - HAMOUR, Christine
 APPLICANT: CORVAT, Nathalie
 APPLICANT: BECK, Alain
 APPLICANT: GOETSCH, Liliane
 TITLE OF INVENTION: MOLECULE OF PHARMACEUTICAL INTEREST COMPRISING AT ITS TERMINAL A GLUTAMIC ACID OR A GLUTAMINE IN THE FORM OF A PHYSIOLOGICALLY ACCEPTABLE STRONG ACID
 FILE REFERENCE: 343 727 - US
 CURRENT APPLICATION NUMBER: US/10/239,313A
 CURRENT FILING DATE: 2002-05-19
 PRIOR APPLICATION NUMBER: FR 00/03711
 PRIOR FILING DATE: 2000-03-23
 PRIOR APPLICATION NUMBER: PCT 01/70772
 PRIOR FILING DATE: 2001-03-22
 NUMBER OF SEQ ID NOS: 697
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 54
 LENGTH: 19
 TYPE: PRT
 ORGANISM: Plasmodium malariae
 US-10-339-313A-54

Query Match 93.7%; Score 89; DB 14; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.3e-07; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIAKMEKASSVFNVNNS 20
 Db 1 EKKIAKMEKASSVFNVNNS 19

RESULT 6
 US-09-932-165-1482
 Sequence 1482, Application US/09932165
 Publication No. US20030134784A1
 GENERAL INFORMATION:
 APPLICANT: RAITANO, ARTHUR
 APPLICANT: CHALLITA-BID, PIA M.
 APPLICANT: SAFFRAN, DOUGLAS
 APPLICANT: AFAR, DANIEL
 APPLICANT: LEVIN, ELANA
 APPLICANT: HUBERT, RENE
 APPLICANT: GE, WANGMAO
 APPLICANT: JAKOBOWITZ, AYA
 TITLE OF INVENTION: NUCLEAR ACIDS AND CORRESPONDING PROTEINS ENTITLED CATHF211 USEFUL IN TREATMENT AND DETECTION OF CANCER
 FILE REFERENCE: 51158-20014.00
 CURRENT APPLICATION NUMBER: US/09/932,165
 PRIOR APPLICATION NUMBER: 2001-08-17
 PRIOR FILING DATE: 2000-08-17
 NUMBER OF SEQ ID NOS: 1508
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 1482
 LENGTH: 21
 TYPE: PRT
 ORGANISM: Plasmodium falciparum
 US-09-932-165-1482

Query Match 93.7%; Score 89; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIAKMEKASSVFNVNNS 20

RESULT 7
 US-09-935-384-710
 Sequence 710, Application US/09935384
 Publication No. US20030166526A1
 GENERAL INFORMATION:
 APPLICANT: CHALLITA-BID, PIA
 APPLICANT: HUBERT, RENE
 APPLICANT: RAITANO, ARTHUR
 APPLICANT: AFAR, DANIEL
 APPLICANT: LEVIN, ELANA
 APPLICANT: PARIS, MARY
 APPLICANT: GE, WANGMAO
 APPLICANT: JAKOBOWITZ, AYA
 TITLE OF INVENTION: NUCLEAR ACID AND CORRESPONDING PROTEIN NAMED 158P1H4
 TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND OTHER CANCERS
 FILE REFERENCE: 51158-20033.00
 CURRENT APPLICATION NUMBER: US/09/935,384
 CURRENT FILING DATE: 2001-08-22
 PRIOR APPLICATION NUMBER: 60/227,098
 PRIOR FILING DATE: 2000-08-22
 NUMBER OF SEQ ID NOS: 743
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 710
 LENGTH: 21
 TYPE: PRT
 ORGANISM: Plasmodium falciparum
 US-09-935-384-710

Query Match 93.7%; Score 89; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIAKMEKASSVFNVNNS 20
 Db 3 EKKIAKMEKASSVFNVNNS 21

RESULT 8
 US-09-942-052-711
 Sequence 711, Application US/09942052
 Publication No. US20030170626A1
 GENERAL INFORMATION:
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Paris, Mary
 APPLICANT: Hubert, Rene S.
 APPLICANT: Afar, Daniel
 APPLICANT: Ge, Wangmao
 APPLICANT: Challita-Bid, Pia M.
 TITLE OF INVENTION: NUCLEAR ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
 TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
 FILE REFERENCE: 51158-20028.00
 CURRENT APPLICATION NUMBER: US/09/942,052
 CURRENT FILING DATE: 2001-08-28
 PRIOR APPLICATION NUMBER: 60/228,432
 PRIOR FILING DATE: 2000-08-28
 NUMBER OF SEQ ID NOS: 744
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 711
 LENGTH: 21
 TYPE: PRT
 ORGANISM: Plasmodium falciparum
 US-09-942-052-711

Query Match 93.7%; Score 89; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIAKMEKASSVFNVNNS 20

Db ||||| 3 EKKIAKNEKASSVFNVNS 21

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/197,484
FILING DATE: 16-FEB-1994

APPLICATION NUMBER: US 07/935,811
FILING DATE: 26-AUG-1992

APPLICATION NUMBER: US 07/874,491
FILING DATE: 27-APR-1992

APPLICATION NUMBER: US 07/627,682
FILING DATE: 29-JAN-1992

APPLICATION NUMBER: US 07/749,568
FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:
NAME: Parmilee, Steven W.

REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14-137-26-4
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 467-9600
TELEFAX: (206) 623-6793

INFORMATION FOR SEQ ID NO: 97;
SEQUENCE CHARACTERISTICS:

CURRENT FILING DATE: 2002-05-20
TYPE: amino acids

SPREADNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: peptide

FEATURE: NAME/KEY: Peptide
LOCATION: 1..21.

OTHER INFORMATION: /note= "Malaria circumsporozoite
37-398"

US-10-128-711-97

Query Match 93.7%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.6e-07; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Delins 0;

Qy 2 EKKIAKNEKASSVFNVNS 20
Db 3 EKKIAKNEKASSVFNVNS 21

RESULT

US-10-116-118-33
Sequence 33, Application US/10116118
Publication No. US20030143672A1

GENERAL INFORMATION:
APPLICANT: Targri, Shabnam
Satte, Alessandro

APPLICANT: Ishioka, Glenn
Fikes, John D.

TITLE OF INVENTION: Heteroclitic Analogs and Related Methods
FILE REFERENCE: 2000-090003

CURRENT APPLICATION NUMBER: US/10/116,118
CURRENT FILING DATE: 2002-08-07

PRIOR APPLICATION NUMBER: US 60/166,529
PRIOR FILING DATE: 1999-11-18

PRIOR APPLICATION NUMBER: US 60/239,008
PRIOR FILING DATE: 2000-10-06

NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 33

LENGTH: 21
TYPE: PRT
ORGANISM: Plasmodium falciparum

US-10-116-118-33
Query Match 93.7%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.6e-07; Indels 0; Gaps 0;

Qy 2 EKKIAKNEKASSVFNVNS 20
Db |||||

RESULT 9
US-10-001-469-1404
; Sequence 1404, Application US/10001469
; Publication No. US200300963A1
; GENERAL INFORMATION:
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: APAR, DANIEL
; APPLICANT: SAFFRAN, DOUGLAS
; APPLICANT: HUBERT, RENE
; APPLICANT: PARIS, MARY
; APPLICANT: CHAILLET, EDD, PIA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED
; FILE NUMBER: 101P3A1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20024-20
; CURRENT APPLICATION NUMBER: US/10/001,469
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/157,902
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/291,118
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 09/680,728
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 2888
; SEQ ID NO: 1404
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-001-469-1404

Query Match 93.7%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.6e-07; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Delins 0;

Qy 2 EKKIAKNEKASSVFNVNS 20
Db 3 EKKIAKNEKASSVFNVNS 21

RESULT 10
US-10-128-711-97
; Sequence 97, Application US/10128711
; Publication No. US200300963A1

GENERAL INFORMATION:
APPLICANT: VITIELLO, Maria A.
CHESNUT, Robert W.
SATTE, Alessandro D.
CHILIS, Esteban
GRAY, Howard

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
CTL IMMUNITY
NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend Khourie and Crea
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/128,711
FILING DATE: 22-Apr-2002
CLASSIFICATION: <Unknown>

Db 3 EKKIAMEKASSVNTVNS 21

RESULT 12
 US-10-062-109A-761
 ; Sequence 761, Application US/10062109A
 ; Publication No. US20030165505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys
 ; APPLICANT: Challita-Eid, Pia M.
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Faris, Mary
 ; APPLICANT: Hubert, Rene S.
 ; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
 ; Entitled 161P2F10B Useful in Treatment and Detection of
 ; Cancer

Query Match 93.7%; Score 89; DB 14; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIAMEKASSVNTVNS 20
 Db 3 EKKIAMEKASSVNTVNS 21

RESULT 13
 US-10-005-480A-761
 ; Sequence 761, Application US/10005480A
 ; Publication No. US20030191073A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys
 ; APPLICANT: Challita-Eid, Pia M.
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Faris, Mary
 ; APPLICANT: Hubert, Rene S.
 ; APPLICANT: Morrison, Karen Jane Meyrick
 ; APPLICANT: Jakobovits, Aya
 ; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
 ; Entitled 161P2F10B Useful in Treatment and Detection of
 ; Cancer

Query Match 93.7%; Score 89; DB 14; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIAMEKASSVNTVNS 20
 Db 3 EKKIAMEKASSVNTVNS 21

RESULT 14
 US-10-277-292-652
 ; Sequence 652, Application US/10277292
 ; Publication No. US20030199470A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Faris, Mary
 ; APPLICANT: Raitano, Arthur
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Levitt, Elana
 ; APPLICANT: Challita-Eid, Pia
 ; APPLICANT: Jakobovitz, Aya
 ; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein Named 161P1D7
 ; Entitled 161P1D7 Useful in the Treatment and Detection of Bladder and
 ; Other Cancers

Query Match 93.7%; Score 89; DB 14; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIAMEKASSVNTVNS 20
 Db 3 EKKIAMEKASSVNTVNS 21

RESULT 15
 US-10-291-241-26
 ; Sequence 26, Application US/10291241
 ; Publication No. US20030206505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys, Inc.
 ; APPLICANT: Ayu Jakobovits
 ; APPLICANT: Arthur B. Raitano
 ; APPLICANT: Mary Paris
 ; APPLICANT: Rene S. Hubert
 ; APPLICANT: Wangmao Ge
 ; APPLICANT: Karen Jane Meyrick Morrison
 ; APPLICANT: Robert Kendall Morrison
 ; APPLICANT: Pia M. Challita-Eid
 ; APPLICANT: Arthur B. Raitano
 ; APPLICANT: Mary Paris
 ; APPLICANT: Rene S. Hubert
 ; APPLICANT: Karen Jane Meyrick Morrison
 ; APPLICANT: Robert Kendall Morrison
 ; APPLICANT: Pia M. Challita-Eid
 ; APPLICANT: Arthur B. Raitano
 ; APPLICANT: Mary Paris
 ; APPLICANT: Rene S. Hubert
 ; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
 ; Entitled 161P2F10B Useful in Treatment and Detection of
 ; Cancer

Query Match 93.7%; Score 89; DB 14; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIAMEKASSVNTVNS 20
 Db 3 EKKIAMEKASSVNTVNS 21

TYPE: PRT
ORGANISM: Plasmodium falciparum
US-10-291-241-26
Query Match 93.7%; Score 89; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 EKKIAKMEKASSYFNVNS 20
| | | | | | | | | | | | | | | | | | | | | |
3 EKKIAKMEKASSYFNVNS 21
Db
Search completed: March 10, 2004, 10:25:47
Job time : 15.7198 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 : Search time 6.22568 Seconds
 (without alignments)
 309.015 Million cell updates/sec

Title: US-09-848-834A-3

Perfect score: 95
 Sequence: 1 DEKKIAKMRKASSVFNVNS 20

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%

Listing First 45 summaries

Database : PIR78:
 1: Pir1:
 2: Pir2:
 3: Pir3:
 4: Pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	85.3	388	A39756	circumsporozoite P
2	81	85.3	405	S05428	circumsporozoite P
3	81	85.3	412	1 OZZQAF	circumsporozoite P
4	81	85.3	424	A54533	circumsporozoite P
5	81	85.3	442	2 A34529	circumsporozoite P
6	55	57.9	264	A44969	circumsporozoite P
7	55	57.9	332	1 OZZQMB	circumsporozoite P
8	55	57.9	348	1 OZZQAK	circumsporozoite P
9	55	57.9	367	1 OZZQMY	circumsporozoite P
10	51	53.7	429	A54504	circumsporozoite P
11	51	53.7	485	2 A60610	circumsporozoite P
12	46	48.4	1534	A56734	circumsporozoite P
13	45	47.4	378	1 OZZQAB	circumsporozoite P
14	45	47.4	378	1 OZZQAL	circumsporozoite P
15	45	47.4	398	1 OZZQAS	circumsporozoite P
16	45	47.4	401	1 OZZQAC	circumsporozoite P
17	45	47.4	419	1 OZZQAM	circumsporozoite P
18	44	46.3	255	2 E81265	probable periplasm
19	44	46.3	441	1 A45345	gag polyprotein -
20	44	46.3	482	1 JC5593	amphiphysin II2 -
21	44	46.3	529	2 A90328	hypothetical prote
22	44	46.3	994	2 S00960	hypothetical prote
23	43.5	45.8	398	2 JC6164	circumsporozoite P
24	43	45.3	280	2 H90556	fructose-bisphosph
25	43	45.3	343	2 A23319	circumsporozoite P
26	43	45.3	367	2 A32068	circumsporozoite P
27	43	45.3	378	2 A45571	circumsporozoite P
28	43	45.3	386	2 D41156	circumsporozoite P
29	43	45.3	387	2	

RESULT 1

A39756

circumsporozoite Protein - Plasmodium reichenowi

C;Species: Plasmodium reichenowi
 C;Date: 14-Feb-1992 #text_change 14-Feb-1992

C;Accession: A39756

R;ai, A.A.; Goldman, I.F.

J. Biol. Chem. 266: 6686-6689, 1991

A;Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria

A;Reference number: A39756; MUID:91201303; PMID:2016283

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-388 <LAL>

A;Cross references: GB:MG6972; PID:9160228; PID:AAA29561.1; PID:9160229

C;Superfamily: circumsporozoite protein, thrombospondin type 1 repeat homology <THRI>

F;312-366/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match %: 85.3%; Score: 81; DB: 2; Length: 388;

Best Local Similarity: 89.5%; Pred. No.: 4.7e-05;

Matches: 17; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

Query Match %: 85.3%; Score: 81; DB: 2; Length: 388;

Best Local Similarity: 89.5%; Pred. No.: 4.7e-05;

Matches: 17; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

RESULT 2

S05428

circumsporozoite Protein - malaria parasite (Plasmodium falciparum)

C;Species: Plasmodium falciparum

C;Accession: S05428; A45527; T60657

R;Campbell, J.R.

Nucleic Acids Res. 17: 5854, 1989

A;Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate v.

A;Reference number: S05428; MUID:89345180; PMID:2668895

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-405 <CAM>

A;Cross references: EMBL:X15363

R;Caspers, P.; Gantz, R.; Matile, H.; Pink, J.R.; Sinigaglia, F.

Mol. Biochem. Parasitol. 35: 190, 1989

A;Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate

A;Reference number: A45527; MUID:89364998; PMID:2671723

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-405 <CAS>

A;Cross references: GB:J04650; PID:91201302; PMID:9160168; PID:AAA29527.1; PID:9160169

R;Lockyer, M.J.; Marsh, K.; Newbold, C.I.

ALIGNMENTS

Mol. Biochem. Parasitol. 37, 275-280, 1989
 A;Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell epitopes
 A;Reference number: A60657; MUID:90114334; PMID:2481827
 A;Accession: I6657
 A;Status: Preliminary; not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 319-336, 354-373 <LOC>
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F;329-383/Domain: thrombospondin type 1 repeat homology <THR1>
 Query Match 85.3%; Score 81; DB 2; Length 405;
 Best Local Similarity 89.5%; Pred. No. 4.9e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 EKKIARMKASSFVNVS 20
 Db 373 EKKIARMKCSSFVNVS 391

RESULT 3
 OZQAF
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate IMTM22)
 C;Species: Plasmodium falciparum
 C;Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jun-2000
 C;Accession: A03188
 R;Dame, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W.T.
 Science 225, 593-599, 1984
 A;Title: Structure of the gene encoding the immunodominant surface antigen on the sporozoite of Plasmodium yoelii
 A;Reference number: A03388; MUID:84250215; PMID:6204383
 A;Accession: A03188
 A;Molecule type: DNA
 A;Residues: 1-412 <DAM>
 A;Cross-references: GB:K02194; NID:gi160160; PIDN:AAA29524.1; PID:gi160161
 A;Experimental source: Clone 7G8
 C;Comment: Residues 1-16 are the probable signal sequence.
 C;Comment: There are 41 copies of a 4-residue repeating unit in the middle domain of the
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F;336-390/Domain: thrombospondin type 1 repeat homology <THR1>
 Query Match 85.3%; Score 81; DB 2; Length 412;
 Best Local Similarity 89.5%; Pred. No. 5e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 EKKIARMKASSFVNVS 20
 Db 380 EKKIARMKCSSFVNVS 398

RESULT 4
 A54533
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain T4, Thailand)
 C;Species: Plasmodium falciparum
 C;Accession: A54533
 R;Diel, Portillo, H.A.; Nussenzweig, R.S.; Ennea, V.
 Mol. Biochem. Parasitol. 24, 189-194, 1987
 A;Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.
 A;Reference number: A54533; MUID:8731205; PMID:3303373
 A;Accession: A54533
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-424
 A;Cross-references: GB:M19752; NID:gi160216; PIDN:AAA29555.1; PID:gi160217
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology <THR1>
 F;348-402/Domain: thrombospondin type 1 repeat homology <THR1>
 Query Match 85.3%; Score 81; DB 2; Length 424;
 Best Local Similarity 89.5%; Pred. No. 5.1e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 EKKIARMKASSFVNVS 20
 Db 392 EKKIARMKCSSFVNVS 410

RESULT 5
 A54529
 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain Wellcome)
 C;Species: Plasmodium falciparum
 C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
 C;Accession: A54529
 R;Lockyer, M.J.; Schwarz, R.T.
 Mol. Biochem. Parasitol. 22, 101-108, 1987
 A;Title: Strain variation in the circumsporozoite protein gene of Plasmodium falciparum
 A;Reference number: A54529; MUID:87115616; PMID:3443671
 A;Accession: A5429
 A;Status: Preliminary; not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-442 <LOC>
 A;Cross-references: GB:MI5505; NID:gi160214; PIDN:AAA29554.1; PID:gi160215
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C;Keywords: circumsporozoite protein; thrombospondin type 1 repeat homology
 F;366-420/Domain: thrombospondin type 1 repeat homology <THR1>
 Query Match 85.3%; Score 81; DB 2; Length 442;
 Best Local Similarity 89.5%; Pred. No. 5.4e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 EKKIARMKASSFVNVS 20
 Db 410 EKKIARMKCSSFVNVS 428

RESULT 6
 A44669
 circumsporozoite protein precursor - Plasmodium yoelii nigeriensis
 C;Species: Plasmodium yoelii nigeriensis
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-1995
 C;Accession: A44669
 R;Colombe-Gould, V.; Ennea, V.
 Mol. Biochem. Parasitol. 43, 51-58, 1990
 A;Title: Plasmodium yoelii nigeriensis circumsporozoite gene structure and its implication
 A;Reference number: A44961; MUID:91148645; PMID:2290446
 A;Accession: A44669
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-264 <COL>
 A;Cross-references: GB:M32350
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F;199-242/Domain: circumsporozoite protein; thrombospondin type 1 repeat homology <THR1>
 Query Match 57.9%; Score 55; DB 2; Length 264;
 Best Local Similarity 47.4%; Pred. No. 0.4;
 Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 EKKIARMKASSFVNVS 20
 Db 232 DTEICRMKCSSIFVNVS 250

RESULT 7
 OZZOMB
 circumsporozoite protein precursor - Plasmodium berghei (strain NK65)
 C;Species: Plasmodium berghei
 C;Date: 30-Sep-1987 #sequence_revision 28-Jul-1995 #text_change 16-Jul-1999
 C;Accession: A44948; A25083; S13446
 R;Lanar, D.E.
 Mol. Biochem. Parasitol. 39, 151-154, 1990
 A;Title: Sequence of the circumsporozoite gene of Plasmodium berghei ANKA clone and NK65
 A;Reference number: A44948; MUID:90156893; PMID:2406593
 A;Accession: A44948
 A;Molecule type: DNA
 A;Residues: 1-322 <LEN>
 A;Cross-references: GB:M28887
 R;Eichinger, D.J.; Arnot, D.E.; Tam, J.P.; Nussenzweig, V.; Ennea, V.

Mol. Cell. Biol. 6, 3965-3972, 1986
 A;Title: Circumsporozoite protein of Plasmodium berghhei: gene cloning and identification
 A;Reference number: A25083; MUID: 87089740; PMID: 2432395
 A;Accession: A25083
 A;Molecule type: DNA
 A;Residues: 'G,B:25445; NID:9160177; PIDN:AAA29531.1';
 A;Cross-references: GB:25445; NID:9160177; PIDN:AAA29531.1; PID:g160178
 C;Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C;Keywords: sporozoite; surface antigen; tandem repeat
 F;1-23/Domain: signal sequence #status predicted <SIG>
 F;24-33/Domain: circumsporozoite protein #status predicted <SIG>
 F;34-43/Region: 8-residue repeats
 F;44-53/Region: 8-residue repeats
 F;258-310/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 57.9%; Score 55; DB 1; Length 332;
 Best Local Similarity 47.4%; Pred. No. 0.51;
 Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 EKKIAKMEKASSYFVNVS 20
 Db 300 DTEICKMDKCSSIFNIVSN 318

RESULT 8

OZZQBK
 circumsporozoite protein precursor - Plasmodium berghhei (strain ANKA clone 2.34L)
 N;Alternate names: sporozoite surface antigen
 C;Species: Plasmodium berghhei
 C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
 C;Accession: S07873; S12571
 R;Lockyer, M.J.; Davies, C.S.; Suhrbier, A.; Sinden, R.E.
 Nucleic Acids Res. 18, 376, 1990
 A;Title: Nucleotide sequence of the Plasmodium berghhei circumsporozoite protein gene from
 A;Reference number: S07873; MUID:9021834; PMID:2183186
 A;Accession: S07873
 A;Molecule type: DNA
 A;Residues: 1-348 <LOC>
 A;Cross-references: EMBL:X17606
 R;Lockyer, M.J.
 submitted to the EMBL Data Library, November 1989

A;Reference number: S12571
 A;Molecule type: DNA
 A;Residues: 1-59, 'I, 61-81, 83-348 <LOC2>;
 A;Cross-references: EMBL:X17606; NID:99784; PID:CAA35608.1'; PID:99785
 C;Keywords: tandem repeat
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-205/Region: 8-residue repeats
 F;215-247/Region: 2-residue repeats
 F;24-336/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 57.9%; Score 55; DB 1; Length 348;
 Best Local Similarity 47.4%; Pred. No. 0.53;
 Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EKKIAKMEKASSYFVNVS 20

Db 316 DTEICKMDKCSSIFNIVSN 334

RESULT 9

OZZQMY
 circumsporozoite protein precursor - Plasmodium yoelii
 N;Alternate names: sporozoite surface antigen
 C;Species: Plasmodium yoelii
 C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
 C;Accession: A26271
 R;Lal, A.A.; de la Cruz, V.F.; Welsh, J.A.; Charoenvit, Y.; Maloy, W.L.; McCutchan, T.
 J. Biol. Chem. 261, 2357-2360, 1987
 A;Title: Structure of the gene encoding the circumsporozoite protein of Plasmodium yoelii
 A;Reference number: A26271; MUID:87137555; PMID:3102479
 A;Accession: A26271
 A;Molecule type: DNA
 A;Residues: 1-367 <LAL>
 A;Cross-references: GB:J02695; NID:9160222; PIDN:AAA29558.1; PID:g160223
 C;Comment: There are three distinct regions in the mature circumsporozoite protein, thobic membrane-anchoring sequence.
 C;Comment: There are 15 copies of a 6-residue repeat and 8 copies of a 4-residue repea
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C;Keywords: sporozoite; surface antigen; tandem repeat
 F;1-19/Domain: signal sequence #status predicted <SIG>
 F;20-367/Product: circumsporozoite protein #status predicted <MAT>
 F;139-226/Region: 6-residue repeats (Q-G-P-G-A-P)
 F;229-260/Region: 4-residue repeats (Q-Q-P-P)
 F;293-345/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 57.9%; Score 55; DB 1; Length 367;
 Best Local Similarity 47.4%; Pred. No. 0.56;
 Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 EKKIAKMEKASSYFVNVS 20
 Db 335 DTEICKMDKCSSIFNIVSN 355

RESULT 10

A54504
 circumsporozoite protein - Plasmodium malariae (strain Uganda 1-CDC)
 C;Species: Plasmodium malariae
 C;Date: 06-Oct-1994 #sequence_revision 04-Nov-1994 #text_change 20-Aug-1999
 C;Accession: A54504
 R;Lal, A.A.; de la Cruz, V.F.; Campbell, G.H.; Procell, P.M.; Collins, W.E.; McCutchan Mol. Biochem. Parasitol. 30, 291-294, 1988
 A;Title: Structure of the circumsporozoite gene of Plasmodium malariae.
 A;Reference number: A54504; MUID:8900027; PMID:3056537
 A;Accession: A54504
 A;Molecule type: DNA
 A;Residues: 1-429 <LAL>
 A;Cross-references: GB:J03992; NID:9160220; PIDN:AAA29557.1; PID:g160221
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C;Keywords: tandem repeat
 F;354-407/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 53.7%; Score 51; DB 2; Length 429;
 Best Local Similarity 47.4%; Pred. No. 2.8;
 Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 EKKIAKMEKASSYFVNVS 20
 Db 397 DTEICKMDKCSSIFNIVSN 415

RESULT 11

A60610
 circumsporozoite protein precursor - Plasmodium brasiliannum
 C;Species: Plasmodium brasiliannum
 C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 20-Aug-1999
 C;Accession: A60610; A28615
 R;di Giovanni, L.; Cochrane, A.H.; Enea, V.
 Exp. Parasitol. 70, 373-381, 1990
 A;Title: On the evolutionary history of the circumsporozoite protein in plasmodia.
 A;Reference number: A60610; MUID:90214818; PMID:2323391

A;Accession: A60610
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-485 <DNA>
A;Experimental source: strain Columbian
R.;Lai, A.; de la Cruz, V.-B.; Collins, W.B.; Campbell, G.H.; Procell, P.M.; McCutchan, J.; Biol. Chem. 263, 5455-5458, 1988
A;Title: Circumsporozoite protein gene from Plasmodium brasiliannum. Animal reservoirs for
A;Reference number: A28615; MUID:8188854; PMID:3128542
A;Accession: A28615
A;Molecule type: DNA
A;Residues: 93-485 <DNA>
A;Cross-references: GB:J03203; NID:9160212; PID:AA29553.1; PID:9160213
C;Keywords: sporozoite; surface antigen; tandem repeat homology
P;1-20/Domain: signal sequence #status predicted <SG>
P;21-485/Domain: circumsporozoite protein; tandem repeat homology
P;41-463/Domain: circumsporozoite protein #status predicted <SG>
P;114-369/Region: 4-residue repeats
P;410-463/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match Score 51; DB 2; Length 485;
Best Local Similarity 53.7%; Pred. No. 3.2%; Indels 0; Gaps 0;
Matches 9; Conservative 6; Mismatches 4;

Qy 2 EKKIAKMEKASSVENVVNS 20
Db 453 ETEVCTMDKAGIFNVVSN 471

RESULT 12
A56734
ribosome receptor, 180k - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Oct-1999
C;Accession: A56734
R.;Wanker, E.E.; Sun, Y.; Savitz, A.J.; Meyer, D.I.
J. Cell Biol. 130, 29-39, 1995
A;Title: Functional characterization of the 180-kD ribosome receptor in vivo.
A;Reference number: A56734; MUID:95310363; PMID:7793375
A;Accession: A56734
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-153 <RNA>
A;Cross-references: GB:J87224; NID:9984113; PID:CAA60676.1; PID:9984114
C;Keywords: endoplasmic reticulum; membrane protein; protein biosynthesis
P;1998-743/Region: 10-residue repeats (N-Q-G-K-A-E-G-A-P)

Query Match Score 46; DB 2; Length 1534;
Best Local Similarity 63.2%; Pred. No. 68; Indels 2; Gaps 1;
Matches 12; Conservative 3; Mismatches 2;

Qy 2 EKKIAKMEKASSVENVVNS 20
Db 143 EKKVAKVEPA-VSSVNVNS 159

RESULT 13
OZZQAB
circumsporozoite protein precursor - Plasmodium cynomolgi (strain Berok)
C;Species: Plasmodium cynomolgi
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
C;Accession: D26255
R.Galinski, M.R.; Arnot, D.E.; Barnwell, J.W.; Nussenzweig, R.S.; Enea, Cell 48, 311-319, 1987
A;Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.
A;Reference number: A90889; MUID:87102878; PMID:3802156
A;Accession: C26255
A;Molecule type: DNA
A;Residues: 1-398 <Gal>
C;Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.

C;Species: Plasmodium cynomolgi
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
C;Accession: D26255
A;Molecule type: DNA
A;Residues: 1-378 <Gal>
C;Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.

C;Species: Plasmodium cynomolgi
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 07-Nov-1997
C;Accession: D26255
A;Molecule type: DNA
A;Residues: 1-378 <Gal>
C;Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.

SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-019-010-2

Query Match 96.3%; Score 52; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.022; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Gaps 0;

Qy 2 HWSYGLRP 9
Db 2 HWSYGLRP 9

RESULT 3
US-09-810-601-1
; Sequence 1, Application US/09810601
; Patent No. US20017545A1
; GENERAL INFORMATION:
; APPLICANT: Donovan, Stephen
; TITLE OF INVENTION: Compositions and Methods for Treating Gonadotrophin
; TITLE OF INVENTION: Related Illnesses
; FILE REFERENCE: 2947
; CURRENT APPLICATION NUMBER: US/09/810,601
; CURRENT FILING DATE: 2001-03-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: OTHER INFORMATION: Description of Artificial Sequence: Gonadotrophin
; OTHER INFORMATION: Release Hormone Analogue
; NAME/KEY: MOD_RES
; LOCATION: (1)
; OTHER INFORMATION: Xaa at position 1 is PyroGlu; PYRROLIDONE
; OTHER INFORMATION: Xaa at position 1 is PyroGlu; PYRROLIDONE
; NAME/KEY: MOD_RES
; LOCATION: (10)
; OTHER INFORMATION: Xaa at position 10 is Gly-NH2; AMIDATION;
; PUBLICATION INFORMATION:
; PATENT DOCUMENT NUMBER: 5611226
; PATENT FILING DATE: 1993-07-07
; PUBLICATION DATE: 1997-05-20
; PUBLICATION INFORMATION:
; AUTHORS: Roussele, Christophe
; JOURNAL: Mol. Pharmacol.
; VOLUME: 57
; PAGES: 679-686
; DATE: 2000
US-09-810-601-2

Query Match 96.3%; Score 52; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.022; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Gaps 0;

Qy 2 HWSYGLRP 9
Db 2 HWSYGLRP 9

RESULT 4
US-09-810-601-2
; Sequence 2, Application US/09810601
; Patent No. US20017545A1
; GENERAL INFORMATION:
; APPLICANT: Donovan, Stephen
; TITLE OF INVENTION: Compositions and Methods for Treating Gonadotrophin
; TITLE OF INVENTION: Related Illnesses
; FILE REFERENCE: 2947
; CURRENT APPLICATION NUMBER: US/09/810,601
; CURRENT FILING DATE: 2001-03-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: OTHER INFORMATION: Description of Artificial Sequence: Gonadotrophin
; OTHER INFORMATION: Release Hormone Analogue
; NAME/KEY: MOD_RES
; LOCATION: (1)
; OTHER INFORMATION: Xaa at position 1 is PyroGlu; PYRROLIDONE
; OTHER INFORMATION: Xaa at position 1 is PyroGlu; PYRROLIDONE
; NAME/KEY: MOD_RES
; LOCATION: (10)
; OTHER INFORMATION: Xaa at position 10 is Gly-NH2; AMIDATION;
; PUBLICATION INFORMATION:
; PATENT DOCUMENT NUMBER: 5611226
; PATENT FILING DATE: 1993-07-07
; PUBLICATION DATE: 1997-05-20
; PUBLICATION INFORMATION:
; AUTHORS: Roussele, Christophe
; JOURNAL: Mol. Pharmacol.
; VOLUME: 57
; PAGES: 679-686
; DATE: 2000
US-09-810-601-2

Query Match 96.3%; Score 52; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.022; Mismatches 0; Indels 0; Gaps 0;

Copyright GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 10, 2004, 09:16:59 ; Search time 7.85992 Seconds
(without alignments)
260.645 Million cell updates/sec

Title: US-09-848-834A-1

Perfect score: 54

Sequence: 1 XHWSYGLRXP 10

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
 1: /cgn2_6/ptodata/2/pubpaas/us07_pubcomb.pep:
 2: /cgn2_6/ptodata/2/pubpaas/pct_new_pub.pep:
 3: /cgn2_6/ptodata/2/pubpaas/us06_pubcomb.pep:
 4: /cgn2_6/ptodata/2/pubpaas/us05_pubcomb.pep:
 5: /cgn2_6/ptodata/2/pubpaas/us07_new_pub.pep:
 6: /cgn2_6/ptodata/2/pubpaas/potus_pubcomb.pep:
 7: /cgn2_6/ptodata/2/pubpaas/us08_new_pub.pep:
 8: /cgn2_6/ptodata/2/pubpaas/us09_pubcomb.pep:
 9: /cgn2_6/ptodata/2/pubpaas/us10_pubcomb.pep:
 10: /cgn2_6/ptodata/2/pubpaas/us095_pubcomb.pep:
 11: /cgn2_6/ptodata/2/pubpaas/us09c_pubcomb.pep:
 12: /cgn2_6/ptodata/2/pubpaas/us09_new_pub.pep:
 13: /cgn2_6/ptodata/2/pubpaas/us10a_pubcomb.pep:
 14: /cgn2_6/ptodata/2/pubpaas/us10b_pubcomb.pep:
 15: /cgn2_6/ptodata/2/pubpaas/us10c_pubcomb.pep:
 16: /cgn2_6/ptodata/2/pubpaas/us10c_new_pub.pep:
 17: /cgn2_6/ptodata/2/pubpaas/us60_new_pub.pep:
 18: /cgn2_6/ptodata/2/pubpaas/us60_pubcomb.pep:
 * Pred. No. is the number of results Predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	96.3	10	9 US-09-019-010-2	Sequence 2, Appli
2	52	96.3	10	9 US-09-848-834A-1	Sequence 1, Appli
3	52	96.3	10	9 US-09-810-601-1	Sequence 1, Appli
4	52	96.3	10	9 US-09-810-601-2	Sequence 2, Appli
5	52	96.3	10	9 US-09-810-601-3	Sequence 3, Appli
6	52	96.3	10	9 US-09-810-601-4	Sequence 4, Appli
7	52	96.3	10	10 US-09-94-201A-28	Sequence 28, Appli
8	52	96.3	10	10 US-09-94-201A-32	Sequence 32, Appli
9	52	96.3	10	10 US-09-924-1	Sequence 1, Appli
10	52	96.3	10	10 US-09-305-924-1	Sequence 9, Appli
11	52	96.3	10	13 US-10-19-331-2	Sequence 1, Appli
12	52	96.3	10	13 US-10-194-125-1	Sequence 1, Appli
13	52	96.3	10	14 US-10-15-553-1	Sequence 1, Appli
14	52	96.3	10	14 US-10-12-483-1	Sequence 1, Appli
15	52	96.3	10	14 US-10-117-364-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-019-010-2 ; Sequence 2, Application US/09019010
; Patent No. US2001001330A1
; GENERAL INFORMATION:
/ APPLICANT: HARLAND, RICHARD
/ ATTORNEY: MANN, JOHN G.
/ APPLICANT: ACRES, STEPHEN D.
/ TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: ROBINS & ASSOCIATES
/ CITY: MENLO PARK
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94025

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/019,010
/ FILING DATE: 05-FEB-1998
/ CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: US 60/036,883
/ FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
/ NAME: MCCRAKEN, THOMAS P.
/ REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0035
TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 325-7812
/ TELEFAX: (650) 325-7823
/ INFORMATION FOR SEQ ID NO: 2:

Key Location/Qualifiers
 PT Misc-difference 1
 PT /label= Pyr
 FT Modified-site 10
 FT /label= Gly-NH2
 XX BE887639-A.
 XX
 PD 24-AUG-1981.
 XX
 PF 22-MAY-1980; 80US-00152241.
 XX
 PR 22-MAY-1980; 80US-00152241.
 XX
 DA (AMHP) AYERST MCKENNA & HARRISON LTD.
 XX
 E1 Auclair C;
 XX
 DR WPI; 1981-66067D/37.
 XX
 Gonadorelin for treatment of benign prostatic hyperplasia - is the
 PT decapeptide Pyr-His-Tre-Ser-Tyr-Dly-Ieu-Arg-Pro-Gly-nh2 or luteinising
 PT hormone liberating hormone.
 XX
 PS Claim 1: Page 7; 2pp; French.

XX The inventors claim a compn. for the redn. or prevention of undesired
 CC prostatic growth in males. The compn. contains a decapeptide
 CC (gonadorelin) (API0097) with an appropriate vehicle or support. The
 CC parenteral admin. for treating e.g. benign prostatic hyperplasia by
 CC daily doses of 0.035-11.0 (pref. 0.080-2.0) mg/kg.
 CC Gonadorelin is the generic name for LH-RH and is described in US385108.
 CC In the example sc. injection of the compn. significantly reduced the wt.
 CC of seminal vesicles and ventral prostate in rats without affecting the
 CC wt. of the testicles. (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX Sequence 10 AA:
 SQ

Query Match 96.3%; Score 52; DB 1; length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0078;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

Search completed: March 10, 2004, 09:12:00
 Job time : 18.0584 secs

CC antihyperlipaemic; antithrombotic and/or anticancer agent or an agent for
 CC treating menopause. The composition of the invention is useful for
 CC treating and preventing diseases associated with angiogenesis II, such as:
 CC circulatory diseases; hypertension; gastrointestinal disorders and
 CC cancer. The present amino acid sequence is used in the exemplification of
 CC the invention.

XX Sequence 9 AA;

Query Match 96.3%; Score 52; DB 6; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

RESULT 11

AAB23840 ID AAE29940 standard; peptide; 9 AA.
 AC
 XX
 AC AAE29940;
 DT 24-FEB-2003 (first entry)

XX Gonadotrophin releasing hormone analogue (GnRH-A) #3.
 DE
 XX Gonadotrophin releasing hormone analogue; neurotoxin; prostate cancer;
 KW endocrine disorder; gonadotrophin related illness; endometrial cancer;
 KW pancreatic cancer; breast cancer; endometriosis; precocious puberty;
 KW GnRH-A; therapy
 XX Unidentified.

XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "Proglutamic acid"
 FT Modified-site 9
 FT /note= "C-terminal ethylamide"

XX WO200274327-A2.

XX PN

XX PD

XX 26-SEP-2002.

XX PP

PP 11-MAR-2002; 2002WO-US007379.

XX PR 15-MAR-2001; 2001US-00810501.

XX PA (ALLR) ALLERGAN SALES INC.

XX PI Donovan S,

XX DR 2003-018772/01.

XX New agent comprising a light chain and a (modified) heavy chain of a
 PT botulinum, butyricum, or tetani toxin, useful for treating a
 CC gonadotrophin related illness, e.g. breast, prostate pancreatic or
 PT endometrial cancer, or endometriosis.
 XX Disclosure; Page 29; 97pp; English.

XX The invention relates to an agent comprising a neurotoxin preferably
 CC botulinum toxin for treating endocrine disorders for e.g. gonadotrophin related
 CC illness. The agent is useful for treating gonadotrophin related
 CC illness e.g. prostate cancer, endometrial cancer, pancreatic cancer,
 CC breast cancer, endometriosis or precocious puberty. It is also useful for
 CC decreasing gonadotrophin secretion in a mammal. The present sequence is
 CC gonadotrophin releasing hormone analogue (GnRH-A).

XX Sequence 9 AA;

Query Match 96.3%; Score 52; DB 6; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ID ABP96021 standard; peptide; 9 AA.
 XX
 AC ABP96021;
 XX
 DT 02-MAY-2003 (first entry)

XX Gonadotrophin releasing hormone agonist peptide Gonadrelin.
 DE
 XX Gonadotrophin releasing hormone agonist; GnRH agonist; cancer;
 KW premenopausal breast cancer; cytostatic.
 XX
 OS Synthetic.

XX Key Location/Qualifiers
 FH Modified-site 9
 FT /note= 'Pro is C-terminally modified with -NHCH2CH3'

XX WO200287616-A1.

XX PN

XX PD 07-NOV-2002.

XX PR 24-APR-2002; 2002WO-JP004071.

XX PR 25-APR-2001; 2001JP-00128032.

XX (TAKE) TAKEDA CHEM IND LTD.

XX PA

XX PI Igari Y, Kusaka M;

XX DR WO2003-148286/14.

XX Agent for preventing post-operative recurrence of premenopausal breast
 PT cancer contains GnRH agonists or antagonists.
 XX Disclosure; Page 7; 39pp; Japanese.

XX The present invention describes an agent containing Gonadotrophin
 CC releasing hormone (GnRH) agonists or antagonists. GnRH agonist and
 CC antagonists have cytostatic activity. The agent can be used for
 CC preventing post-operative recurrence of premenopausal breast cancer. The
 CC agents have no serious side effects with sustained drug effect over a
 CC long period without frequent administration. The present sequence
 CC represents an example of a GnRH agonist, from the present invention
 XX Sequence 9 AA;

Query Match 96.3%; Score 52; DB 6; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Qy 2 HWSYGLRP 9
 DB 2 HWSYGLRP 9

RESULT 13
 AP10416 ID API0416 standard; peptide; 10 AA.
 XX AC API0416;
 XX DT 25-MAR-2003 (revised)
 DT 10-MAR-2003 (revised)

PT Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;
 XX WPI; 2001-112059/12.
 XX Modifying and attaching therapeutic peptides to albumin prevents peptide degradation, useful for increasing length of in vivo activity.

PT Disclosure; Page 238; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I) comprising a therapeutically active amino acid region (II) and a reactive group (II) (e.g. succinimidyl and maleimido groups) attached to a less therapeutically active amino acid region (IV), which covalently bonds with amino/hydroxy/ol groups on blood components to form a peptide stabilized therapeutic peptide composed of 3-50 amino acids. (I) are useful for modifying therapeutic peptides e.g. hormones, growth factors and neurotransmitters, to protect them from peptidase activity in vivo for the treatment of various disorders. Endogenous therapeutic peptides are not suitable as drug candidates as they require frequent administration due to rapid degradation by peptidases in the body. Modifying and attaching therapeutic peptides to albumin prevents or reduces the action of peptidases to increase length of activity (half life) and specificity as bonding to large molecules decreases intracellular uptake and interference with physiological processes.

CC AAB9229 to AAB9241 represent Peptides which can be used in the exemplification of the present invention

XX Sequence 9 AA;

Query Match 96.3%; Score 52; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 AC AAB9236;

QY 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

RESULT 9
 AAB9236 standard; peptide; 9 AA.

AC AAB9236;

DT 26-MAR-2001 (first entry)

DE GnRH peptide.

FW GnRH-III; autoimmune disease; transplant rejection; retroviral disease;

KW graft-versus-host disease; lymphoproliferative disease;

KW gonadotropin-releasing hormone.

XX Petromyzon marinus.

Key Modified-site 1 Location/Qualifiers

FT /note= "Lined to Glucagon-like peptide"

XX WO2000074724-A2.

XX PD 14-DEC-2000.

PP 05-JUN-2000; 2000WO-GB002014.

XX PR 03-JUN-1999; 99GB-00012807.

PR 03-JUN-1999; 99US-013752P.

XX PA (BIOI-) BIO INNOVATION LTD.

XX PI Franks CR, Della Bitta R, Maitland NJ, Knight DJ;

XX DR WPI; 2001-061658/07.

XX Novel product comprising proliferatively active moiety linked to genetic material, useful as vectors for protected nucleic acid material and as mitogen to stimulate proliferation of target cell.

XX Disclosure; Page 4; 49pp; English.

XX The present invention relates to a product comprising a proliferatively active moiety (PAM) linked to nucleic acid material which is associated with a protective material. The PAM product is useful for manufacturing a medicament for treating e.g. an autoimmune disease, transplant rejection, retroviral disease, graft-versus-host-disease, or lymphoproliferative disease, comprising cells bearing a high affinity receptor for PAM. The present sequence is a peptide of Gonadotropin-releasing hormone (GnRH). GnRH is a peptide hormone, which has high affinity receptors, and therefore can be used in the present invention

XX SQ Sequence 9 AA;

Query Match 96.3%; Score 52; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 Db 1 HWSYGLRP 8

RESULT 10

AA01785
 ID AA01785 standard; peptide; 9 AA.
 XX AC AA01785;
 XX DT 04-AUG-2003 (first entry)
 XX DE Sustained release composition-related Peptide #5.
 XX KW Sustained release composition; angiotensin II antagonist; menopause; angiotensin II-related disease: circulatory disease; hypertension; gastrointestinal disorder; cancer.

XX OS Unidentified.

XX Key FH Location/Qualifiers
 FF Modified-site 1
 FT /note= "Pyro glutamic acid"
 FT Modified-site 9
 FT /note= "Modified with NH-CH3"

XX WO2003013609-A1.

XX PN 20-FBB-2003.
 XX PD 01-AUG-2002; 2002WO-JP007862.
 XX PR 03-AUG-2001; 2001JP-00236794.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX DR (KAWA.) KAWAMURA R.
 XX PI Kusumoto K, Hoshino T;
 XX DR WPI; 2003-256500/25.

XX Sustained release composition comprises angiotensin II antagonist and antihypertensive, antipolyglycemic, antihyperlipemic, antithrombotic or anticancer agent or agent for treating menopause.

XX PS Disclosure; Page 37; 120pp; Japanese.

XX The invention comprises sustained release compositions that contain:
 CC angiotensin II antagonist; antihypertensive; antihypoglycaemic;
 CC

Best Local Similarity 100.0%; Pred. No. 1.4e+06; Mismatches 0; Indels 0; Gaps 0; Qy 2 HWSYGLRP 9
Matches 8; Conservative 0; N mismatches 0; DB 2 HWSYGLRP 9

RESULT 4
AAB15363 standard; peptide; 9 AA.
ID AAW94891
XX
AC AAW94891;
XX DT 11-MAY-1999 (first entry)
DE LHRH peptide fragment.
XX KW LHRH; immune response; luteinising hormone releasing hormone; DT;
KW diphtheria toxoid; castration; oestrus cycling; aggression; breast;
KW sexual activity; organolectic; livestock; cell growth; malignant;
KW prostate; ovarian; oncofoetal; hyperplastic; pregnancy; endometriosis;
KW inflammatory response.
OS Homo sapiens.
XX PN WO9002180-A1.
XX PR 21-JAN-1999.
XX PF 09-JUL-1998; 98WO-AU000532.
XX PR 09-JUL-1997; 97AU-00007768.
XX DR (CSIC) CSL LTD.
XX PI McNamara MK;
XX DR WPI: 1999-120511/10.
XX
New immunogenic leutinising hormone releasing hormone compositions -
PT comprise LHRH conjugated to diphtheria toxoid and adsorbed to an ionic
PT polysaccharide, used to inhibit reproductive function in animals.
XX Example 3; Page 30; 41PP; English.
XX The invention relates immunogenic composition for eliciting an immune
CC response to luteinising hormone releasing hormone (LHRH). The composition
CC comprises a LHRH-diphtheria toxoid (DT) conjugate adsorbed to an ionic
CC polysaccharide. The LHRH-DT compositions can be used for eliciting an
CC immune response to LHRH, for castrating an animal, for regulating oestrus
CC cycling in a female animal or for inhibiting characteristics induced by
CC the sexual maturation of an animal, e.g., aggression or sexual activity.
CC They can also be used for achieving production gains in livestock, e.g.,
CC reduction or elimination of unwanted organoleptic characteristics from
CC the meat of livestock. They can also be used for inhibiting the growth of
CC breast cells, malignant prostate cells, malignant ovarian cells,
CC malignant oncofoetal cells or hyperplastic cells. They can also be used for
CC down-regulating the libido of an animal. They can also be used for
CC inhibiting pregnancy, prostate enlargement, endometriosis or inflammatory
CC responses. The LHRH compositions induce a more effective immune response
CC against LHRH than the LHRH-carrier-adjuvant compositions. The effective
CC immune response against LHRH results in prevention of the release of the
CC hormones LH and FSH from the anterior pituitary. Sequences AAW94890-93
XX Sequence 9 AA;

Query Match 96.3%; Score 52; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06; Mismatches 0; Indels 0; Gaps 0; SQ

Qy 2 HWSYGLRP 9
DB 1 HWSYGLRP 8

RESULT 5
AAB15363 standard; peptide; 9 AA.
ID AAB15363;
XX AC AAB15363;
XX DT 17-JAN-2001 (first entry)
XX DE Human LHRH peptide SEQ ID NO: 2.
XX Human: LHRH; GnRH; luteinising hormone releasing hormone;
KW Gonadotrophin releasing hormone; fertility control; cancer;
KW prostate; prostate enlargement.
XX OS Homo sapiens.
XX PN WO20001720-A1.
XX PD 20-JUL-2000.
XX PF 24-DEC-1999; 99WO-AU001167.
XX PR 08-JAN-1999; 99AU-00008073.
XX PA (CSIC) CSL LTD.
XX PI Walker J;
XX DR WPI: 2000-475954/41.
XX
Adjuvant composition for manufacturing an immunogenic composition that
PT can elicit an immune response in an animal, comprises an ionic
PT polysaccharide component and a saponin component that is an
PT immunostimulating complex.
XX Disclosure; Page 50; 53pp; English.
XX
The present sequence is a peptide fragment of human luteinising hormone
CC releasing hormone (also known as LHRH, GnRH and gonadotrophin releasing
CC hormone). It was used to demonstrate the novel adjuvant of the invention,
CC which has lower reactivity than previous compositions. Vaccination of
CC humans and animals against LHRH can be used as a method of fertility
CC control, as well as enabling the control and treatment of disorders of
CC the reproductive organs, such as testicular, breast, prostate and ovarian
CC cancers, prostate enlargement and endometriosis. The composition of the
CC invention contains an anionic macromolecule and a saponin component,
CC latter of which is an immunostimulant, and it can also be used with other
CC immunogens including soluble protein antigens, peptide haptens conjugated
CC to a carrier protein and whole viruses
XX Sequence 9 AA;

Query Match 96.3%; Score 52; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06; Mismatches 0; Indels 0; Gaps 0; SQ

Qy 2 HWSYGLRP 9
DB 1 HWSYGLRP 8

RESULT 6
AAB08104 standard; peptide; 9 AA.
ID AAB08104
XX AC AAB08104;
XX DT 04-DEC-2000 (first entry)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 10, 2004, 08:58:48 ; Search time 15.0584 Seconds

(without alignments)

187.635 Million cell updates/sec

Title: US-09-848-834A-1

Perfect score: 54

Sequence: 1 XHWSYGLRXP 10

Scoring table: BLOSSOM62

Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04;*

1: GeneseqP1980s;*

2: GeneseqP1990s;*

3: GeneseqP2000s;*

4: GeneseqP2001s;*

5: GeneseqP2002s;*

6: GeneseqP2003as;*

7: GeneseqP2004s;*

8: GeneseqP2004s;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	96.3	8	ABP96550	Abp96550 Gonadotro
2	52	96.3	9	AAP10414	AAP10414 Luteinisi
3	52	96.3	9	APP50568	App50568 Sequence
4	52	96.3	9	AAW94891	AAW94891 LHR Pept
5	52	96.3	9	AAB15363	Aab15363 Human LHR
6	52	96.3	9	AAB08104	Aab08104 Amino aci
7	52	96.3	9	AAB9079	Aab9079 Luteinisi
8	52	96.3	9	AAB59836	Aab59836 GHRH pept
9	52	96.3	9	AAC16785	Aac16785 Substained
10	52	96.3	9	AAB29840	Aab29840 Gonadotro
11	52	96.3	9	ABP96021	Abp96021 Gonadotro
12	52	96.3	9	AAP10416	Aap10416 Luteinisi
13	52	96.3	10	AAP10411	Aap10411 Luteinisi
14	52	96.3	10	AAP10411	Aap10411 Luteinisi
15	52	96.3	10	AAP10097	Aap10097 Sequence
16	52	96.3	10	AAP20277	Aap20277 Modified
17	52	96.3	10	APP50522	App50522 Gonadotro
18	52	96.3	10	AAP60127	Aap60127 Gonadolib
19	52	96.3	10	AAP61403	Aap61403 Gonadotro
20	52	96.3	10	AAP60576	Aap60576 Novel dec
21	52	96.3	10	AAP70922	Aap70922 Luteinisi
22	52	96.3	10	AAP90530	Aap90530 Sequence
23	52	96.3	10	AAR15713	Aar15713 Peptide #
24	52	96.3	10	AAP26819	Aap26819 Ld releas
25	52	96.3	10	AAR33434	Aar33434 Therapeut

ALIGNMENTS

RESULT 1
ID ABP96550 Standard; peptide; 8 AA.
XX
XX
AC
ABP96550;
XX
DT 27-MAY-2003 (first entry)
DE Gonadotrophin releasing hormone agonist gonadorelin peptide.
XX
KW Gonadotrophin releasing hormone agonist; GnRH agonist; combination drug;
KW Pharmaceutical; breast cancer; endometriosis; myometrium tumour;
KW Alzheimer's disease; circulatory system disorder; menopausal disorder;
KW irregular period; cancer metastasis; premenstrual syndrome; osteopathy;
KW muscular distress; calcium/phosphorus imbalance; SERM; gynaecological;
KW selective oestrogen receptor modulator; cytotoxic; nootropic; muscular;
KW neuroprotective; cardiovascular; endocrine; osteopathic; prostabomegaly;
KW prostate cancer.
XX
Synthetic.
OS
XX
PN WO2003015820-A1.
XX
PD 27-FEB-2003.
XX
PF 08-AUG-2002; 2002WO-JP008130.
XX
PR 10-AUG-2001; 2001JP-00244616.
PA (TAKEDA) TAKEDA CHEM IND LTD.
XX
PI Furuya S, Kubaka M;
XX
DR WPI; 2003-300573/29.
XX
PT Pharmaceutical composition e.g. for breast cancer comprises gonadotrophin
PT releasing hormone agonist and selective estrogen receptor modulator.
XX
PS Disclosure; Page 9; 73pp; Japanese.

The present invention describes a pharmaceutical composition (I) for treating breast cancer, endometriosis, myometrium tumour, Alzheimer's disease, circulatory system disorders, menopausal disorders, irregular periods, cancer metastasis, premenstrual syndrome, muscular distress or osteopathies due to calcium/phosphorus imbalance. (I) comprises a Gonadotrophin releasing hormone (GHRH) agonist and a selective oestrogen receptor modulator (SERM). (I) has cytotoxic, gynaecological, nootropic, neuroprotective, cardiovascular, endocrine, muscular and osteopathic

CC

	Matches	6;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
Qy	2	HWSYGLRP	9							
Db	66	HWSYALSP	73							

RESULT 15

	Q86DB8	PRELIMINARY;	PRT;	165 AA.
ID	Q86DB8;			
AC	Q86DB8;			
DT	01-JUN-2003	(TREMBLrel. 24, Created)		
DT	01-JUN-2003	(TREMBLrel. 24, Last sequence update)		
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)		
DE	Preprogonadotropin-releasing hormone 2 precursor.			
OS	Ciona intestinalis.			
OC	Bukaryotes; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;			
OC	Phlebobranchia; Cionidae; Ciona.			
OX	NCBI_TAXID:7719;			
RN	[1]			
SEQUENCE FROM N.A.				
RX	MEDLINE=2258375; PubMed=12637698;			
RA	Adams .A., Tello J.A., Brinchey J.J., Warby C., Hong D.J.,			
RA	Akinsanya K.O., Mackie G.O., Vale W., Rivier J.E., Sherwood N.M.;			
RT	"Six Novel Gonadotropin-Releasing Hormones Are Encoded as Triplets on			
RT	Each of Two Genes in the Protocordate, Ciona intestinalis.";			
RL	Endocrinology 144:1907-1919 (2003).			
DR	EMBL; AY04708; AA06795.1;			
DR	AY04708;			
DR	GO; GO:0005576; C:extracellular; IEA.			
DR	GO; GO:0005179; F:hormone activity; IEA.			
DR	GO; GO:0007275; P:development; IEA.			
DR	InterPro: IPR002012; GnRH.			
DR	Pfam: PF00446; GnRH_1.			
DR	PROSITE; PS00473; GnRH_3.			
KW	Signal.			
FT	SIGNAL	1	63	Potential.
FT	CHAIN	66	75	gonadotropin-releasing hormone 7.
FT	CHAIN	79	88	gonadotropin-releasing hormone 8.
FT	CHAIN	92	101	gonadotropin-releasing hormone 4.
FT	CHAIN	107		GnRH-associated peptide.
SQ	SEQUENCE	165 AA;	19250 MW;	CB3F748849B33601 CRC64;
Query Match		74.1%	Score 40; DB 5; Length 165;	
Best Local Similarity		75.0%	Pred. No. 14;	
Matches		6;	Mismatches 0;	
			Indels 2;	
			Gaps 0;	

	Q86DB8	PRELIMINARY;	PRT;	165 AA.
ID	Q86DB8;			
AC	Q86DB8;			
DT	01-JUN-2003	(TREMBLrel. 24, Created)		
DT	01-JUN-2003	(TREMBLrel. 24, Last sequence update)		
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)		
DE	Preprogonadotropin-releasing hormone 2 precursor.			
OS	Ciona intestinalis.			
OC	Bukaryotes; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;			
OC	Phlebobranchia; Cionidae; Ciona.			
OX	NCBI_TAXID:7719;			
RN	[1]			
SEQUENCE FROM N.A.				
RX	MEDLINE=2258375; PubMed=12637698;			
RA	Adams .A., Tello J.A., Brinchey J.J., Warby C., Hong D.J.,			
RA	Akinsanya K.O., Mackie G.O., Vale W., Rivier J.E., Sherwood N.M.;			
RT	"Six Novel Gonadotropin-Releasing Hormones Are Encoded as Triplets on			
RT	Each of Two Genes in the Protocordate, Ciona intestinalis.";			
RL	Endocrinology 144:1907-1919 (2003).			
DR	EMBL; AY04708; AA06795.1;			
DR	AY04708;			
DR	GO; GO:0005576; C:extracellular; IEA.			
DR	GO; GO:0005179; F:hormone activity; IEA.			
DR	GO; GO:0007275; P:development; IEA.			
DR	InterPro: IPR002012; GnRH.			
DR	Pfam: PF00446; GnRH_1.			
DR	PROSITE; PS00473; GnRH_3.			
KW	Signal.			
FT	SIGNAL	1	63	Potential.
FT	CHAIN	66	75	gonadotropin-releasing hormone 7.
FT	CHAIN	79	88	gonadotropin-releasing hormone 8.
FT	CHAIN	92	101	gonadotropin-releasing hormone 4.
FT	CHAIN	107		GnRH-associated peptide.
SQ	SEQUENCE	165 AA;	19250 MW;	CB3F748849B33601 CRC64;
Query Match		74.1%	Score 40; DB 5; Length 165;	
Best Local Similarity		75.0%	Pred. No. 14;	
Matches		6;	Mismatches 0;	
			Indels 2;	
			Gaps 0;	

Search completed: March 10, 2004, 09:25:19
Job time : 11.8833 secs

RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AY320285; AAP8407.1; -.
 PT NON_TER
 SQ SEQUENCE 72 AA; 8064 MW; OFACAA422AB792BF CRC64;
 Query Match Score 42; DB 13; Length 72;
 Best Local Similarity 75.0%; Pred. No. 2,3;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 DR 2 HWSYGLRP 9
 DR 1|||:|||
 DR 2 HWSFGLSP 9

RESULT 12
 Q8TH60 PRELIMINARY; PRT; 86 AA.
 AC Q8TH60;
 DT 01-OCT-2002 (TREMBUREL. 22, Last sequence update)
 DT 01-OCT-2002 (TREMBUREL. 24, Last annotation update)
 DT 01-JUN-2003 (TREMBUREL. 24, Last annotation update)
 DE Gonadotropin-releasing hormone.
 OS Alosa sagidissima (American shad).
 BUkaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;
 OC Alosa.
 NCBI_TaxID=34773;
 RN [1]
 SEQUENCE FROM N.A.
 RA Abzrahem E., Gotchiff Y., Zohar Y.;
 RT "American shad (Alosa sagidissima) hrghRH sequence.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AY36381; AAP04492.1;
 GO; GO:0005176; C:extracellular; IEA.
 DR GO; GO:0005179; F:hormone activity; IEA.
 DR InterPro; IPR002725; P:development; IEA.
 DR Pfam; PF00446; GnRH. 1.
 DR PROSITE; PS00447; GnRH. 1.
 SQ SEQUENCE 86 AA; 9554 MW; 8E4921F3CE2350B3 CRC64;
 Query Match Score 41; DB 13; Length 86;
 Best Local Similarity 75.0%; Pred. No. 4,3;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 DR 2 HWSYGLRP 9
 DR 1|||:|||
 DR 24 HWSHGLSP 31

RESULT 13
 Q8GB15 PRELIMINARY; PRT; 1127 AA.
 ID Q8GB15;
 AC Q8GB15;
 DT 01-MAR-2003 (TREMBUREL. 23, Created)
 DT 01-OCT-2003 (TREMBUREL. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBUREL. 25, Last annotation update)
 DB Carbamoyl-phosphate synthase large chain.
 CARB OR BL0068.
 OS Bifidobacterium longum.
 CC Bacteria; Actinobacteria; Actinomycetales; Bifidobacteriales;
 CC Bifidobacteriaceae; Bifidobacterium.
 CX NCBI_TaxID=216816;
 RN [1]
 SEQUENCE FROM N.A.
 STRAIN=NCC 2705;
 MEDLINE=22294977; PubMed=112381787;
 RA Schell M.A., Karimirzou M.; Snel B., Vilanova D., Berger B.,
 RA Pessi G., Zwahlen M.-C., Desire F., Bork P., Delley M.,
 RA Pridmore R.D., Arigoni F.;
 RA "The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
 RL

DR EMBL; AE014621; AAN23934.1; -.
 GO; GO:0005737; C:cytosol; IEA.
 DR EMBL; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0004086; Picardamoyl-phosphate synthase activity; IEA.
 DR GO; GO:0016871; Filaggrin activity; IEA.
 DR GO; GO:0006526; Paragine biosynthesis; IEA.
 DR GO; GO:0006801; Nitrogen metabolism; IEA.
 DR GO; GO:0019855; Pyridoxal base biosynthesis; IEA.
 DR InterPro; IPR06225; Cara L_glu.
 DR InterPro; IPR05483; CPSase_L.
 DR InterPro; IPR05479; CPSase_L_D2.
 DR InterPro; IPR05480; CPSase_L_D3.
 DR InterPro; IPR05481; CPSase_L_N.
 DR InterPro; IPR04362; MGS_11ke.
 DR Pfam; PF0289; CPSase_L_Chain; 2.
 DR Pfam; PF02786; CPSase_L_D2; 3.
 DR Pfam; PF02787; CPSase_L_D3; 1.
 DR Pfam; PF0142; MGS; 1.
 DR PRINTS; PR00038; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00666; CPSASE_1; 2.
 DR PROSITE; PS00677; CPSASE_2; 2.
 KW Complete proteome.
 SQ SEQUENCE 1127 AA; 123194 MW; 2B3C0B4FB3282489 CRC64;
 Query Match Score 41; DB 16; Length 1127;
 Best Local Similarity 85.7%; Pred. No. 74;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 14
 Q86D87 PRELIMINARY; PRT; 164 AA.
 ID Q86D87;
 AC Q86D87;
 DT 01-JUN-2003 (TREMBUREL. 24, Created)
 DT 01-JUN-2003 (TREMBUREL. 24, Last sequence update)
 DR PROSITE; PS00446; GnRH. 1.
 DR PROSITE; PS00447; GnRH. 1.
 DR PROSITE; PS00448; GnRH. 1.
 DR PROSITE; PS00449; GnRH. 1.
 DR PROSITE; PS00450; GnRH. 1.
 DR PROSITE; PS00451; GnRH. 1.
 DR PROSITE; PS00452; GnRH. 1.
 DR PROSITE; PS00453; GnRH. 1.
 DR PROSITE; PS00454; GnRH. 1.
 DR PROSITE; PS00455; GnRH. 1.
 DR PROSITE; PS00456; GnRH. 1.
 DR PROSITE; PS00457; GnRH. 1.
 DR PROSITE; PS00458; GnRH. 1.
 DR PROSITE; PS00459; GnRH. 1.
 DR PROSITE; PS00460; GnRH. 1.
 DR PROSITE; PS00461; GnRH. 1.
 DR PROSITE; PS00462; GnRH. 1.
 DR PROSITE; PS00463; GnRH. 1.
 DR PROSITE; PS00464; GnRH. 1.
 DR PROSITE; PS00465; GnRH. 1.
 DR PROSITE; PS00466; GnRH. 1.
 DR PROSITE; PS00467; GnRH. 1.
 DR PROSITE; PS00468; GnRH. 1.
 DR PROSITE; PS00469; GnRH. 1.
 DR PROSITE; PS00470; GnRH. 1.
 DR PROSITE; PS00471; GnRH. 1.
 DR PROSITE; PS00472; GnRH. 1.
 DR PROSITE; PS00473; GnRH. 1.
 DR PROSITE; PS00474; GnRH. 1.
 DR PROSITE; PS00475; GnRH. 1.
 DR PROSITE; PS00476; GnRH. 1.
 DR PROSITE; PS00477; GnRH. 1.
 DR PROSITE; PS00478; GnRH. 1.
 DR PROSITE; PS00479; GnRH. 1.
 DR PROSITE; PS00480; GnRH. 1.
 DR PROSITE; PS00481; GnRH. 1.
 DR PROSITE; PS00482; GnRH. 1.
 DR PROSITE; PS00483; GnRH. 1.
 DR PROSITE; PS00484; GnRH. 1.
 DR PROSITE; PS00485; GnRH. 1.
 DR PROSITE; PS00486; GnRH. 1.
 DR PROSITE; PS00487; GnRH. 1.
 DR PROSITE; PS00488; CPSASE.
 DR PRINTS; PR00038; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00489; CPSASE_1; 2.
 DR PROSITE; PS00490; CPSASE_2; 2.
 DR PROSITE; PS00491; CPSASE_3; 2.
 DR PROSITE; PS00492; CPSASE_4; 2.
 DR PROSITE; PS00493; CPSASE_5; 2.
 DR PROSITE; PS00494; CPSASE_6; 2.
 DR PROSITE; PS00495; CPSASE_7; 2.
 DR PROSITE; PS00496; CPSASE_8; 2.
 DR PROSITE; PS00497; CPSASE_9; 2.
 DR PROSITE; PS00498; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00499; CPSASE_1; 2.
 DR PROSITE; PS00500; CPSASE_2; 2.
 DR PROSITE; PS00501; CPSASE_3; 2.
 DR PROSITE; PS00502; CPSASE_4; 2.
 DR PROSITE; PS00503; CPSASE_5; 2.
 DR PROSITE; PS00504; CPSASE_6; 2.
 DR PROSITE; PS00505; CPSASE_7; 2.
 DR PROSITE; PS00506; CPSASE_8; 2.
 DR PROSITE; PS00507; CPSASE_9; 2.
 DR PROSITE; PS00508; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00509; CPSASE_1; 2.
 DR PROSITE; PS00510; CPSASE_2; 2.
 DR PROSITE; PS00511; CPSASE_3; 2.
 DR PROSITE; PS00512; CPSASE_4; 2.
 DR PROSITE; PS00513; CPSASE_5; 2.
 DR PROSITE; PS00514; CPSASE_6; 2.
 DR PROSITE; PS00515; CPSASE_7; 2.
 DR PROSITE; PS00516; CPSASE_8; 2.
 DR PROSITE; PS00517; CPSASE_9; 2.
 DR PROSITE; PS00518; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00519; CPSASE_1; 2.
 DR PROSITE; PS00520; CPSASE_2; 2.
 DR PROSITE; PS00521; CPSASE_3; 2.
 DR PROSITE; PS00522; CPSASE_4; 2.
 DR PROSITE; PS00523; CPSASE_5; 2.
 DR PROSITE; PS00524; CPSASE_6; 2.
 DR PROSITE; PS00525; CPSASE_7; 2.
 DR PROSITE; PS00526; CPSASE_8; 2.
 DR PROSITE; PS00527; CPSASE_9; 2.
 DR PROSITE; PS00528; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00529; CPSASE_1; 2.
 DR PROSITE; PS00530; CPSASE_2; 2.
 DR PROSITE; PS00531; CPSASE_3; 2.
 DR PROSITE; PS00532; CPSASE_4; 2.
 DR PROSITE; PS00533; CPSASE_5; 2.
 DR PROSITE; PS00534; CPSASE_6; 2.
 DR PROSITE; PS00535; CPSASE_7; 2.
 DR PROSITE; PS00536; CPSASE_8; 2.
 DR PROSITE; PS00537; CPSASE_9; 2.
 DR PROSITE; PS00538; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00539; CPSASE_1; 2.
 DR PROSITE; PS00540; CPSASE_2; 2.
 DR PROSITE; PS00541; CPSASE_3; 2.
 DR PROSITE; PS00542; CPSASE_4; 2.
 DR PROSITE; PS00543; CPSASE_5; 2.
 DR PROSITE; PS00544; CPSASE_6; 2.
 DR PROSITE; PS00545; CPSASE_7; 2.
 DR PROSITE; PS00546; CPSASE_8; 2.
 DR PROSITE; PS00547; CPSASE_9; 2.
 DR PROSITE; PS00548; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00549; CPSASE_1; 2.
 DR PROSITE; PS00550; CPSASE_2; 2.
 DR PROSITE; PS00551; CPSASE_3; 2.
 DR PROSITE; PS00552; CPSASE_4; 2.
 DR PROSITE; PS00553; CPSASE_5; 2.
 DR PROSITE; PS00554; CPSASE_6; 2.
 DR PROSITE; PS00555; CPSASE_7; 2.
 DR PROSITE; PS00556; CPSASE_8; 2.
 DR PROSITE; PS00557; CPSASE_9; 2.
 DR PROSITE; PS00558; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00559; CPSASE_1; 2.
 DR PROSITE; PS00560; CPSASE_2; 2.
 DR PROSITE; PS00561; CPSASE_3; 2.
 DR PROSITE; PS00562; CPSASE_4; 2.
 DR PROSITE; PS00563; CPSASE_5; 2.
 DR PROSITE; PS00564; CPSASE_6; 2.
 DR PROSITE; PS00565; CPSASE_7; 2.
 DR PROSITE; PS00566; CPSASE_8; 2.
 DR PROSITE; PS00567; CPSASE_9; 2.
 DR PROSITE; PS00568; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00569; CPSASE_1; 2.
 DR PROSITE; PS00570; CPSASE_2; 2.
 DR PROSITE; PS00571; CPSASE_3; 2.
 DR PROSITE; PS00572; CPSASE_4; 2.
 DR PROSITE; PS00573; CPSASE_5; 2.
 DR PROSITE; PS00574; CPSASE_6; 2.
 DR PROSITE; PS00575; CPSASE_7; 2.
 DR PROSITE; PS00576; CPSASE_8; 2.
 DR PROSITE; PS00577; CPSASE_9; 2.
 DR PROSITE; PS00578; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00579; CPSASE_1; 2.
 DR PROSITE; PS00580; CPSASE_2; 2.
 DR PROSITE; PS00581; CPSASE_3; 2.
 DR PROSITE; PS00582; CPSASE_4; 2.
 DR PROSITE; PS00583; CPSASE_5; 2.
 DR PROSITE; PS00584; CPSASE_6; 2.
 DR PROSITE; PS00585; CPSASE_7; 2.
 DR PROSITE; PS00586; CPSASE_8; 2.
 DR PROSITE; PS00587; CPSASE_9; 2.
 DR PROSITE; PS00588; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00589; CPSASE_1; 2.
 DR PROSITE; PS00590; CPSASE_2; 2.
 DR PROSITE; PS00591; CPSASE_3; 2.
 DR PROSITE; PS00592; CPSASE_4; 2.
 DR PROSITE; PS00593; CPSASE_5; 2.
 DR PROSITE; PS00594; CPSASE_6; 2.
 DR PROSITE; PS00595; CPSASE_7; 2.
 DR PROSITE; PS00596; CPSASE_8; 2.
 DR PROSITE; PS00597; CPSASE_9; 2.
 DR PROSITE; PS00598; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00599; CPSASE_1; 2.
 DR PROSITE; PS00600; CPSASE_2; 2.
 DR PROSITE; PS00601; CPSASE_3; 2.
 DR PROSITE; PS00602; CPSASE_4; 2.
 DR PROSITE; PS00603; CPSASE_5; 2.
 DR PROSITE; PS00604; CPSASE_6; 2.
 DR PROSITE; PS00605; CPSASE_7; 2.
 DR PROSITE; PS00606; CPSASE_8; 2.
 DR PROSITE; PS00607; CPSASE_9; 2.
 DR PROSITE; PS00608; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00609; CPSASE_1; 2.
 DR PROSITE; PS00610; CPSASE_2; 2.
 DR PROSITE; PS00611; CPSASE_3; 2.
 DR PROSITE; PS00612; CPSASE_4; 2.
 DR PROSITE; PS00613; CPSASE_5; 2.
 DR PROSITE; PS00614; CPSASE_6; 2.
 DR PROSITE; PS00615; CPSASE_7; 2.
 DR PROSITE; PS00616; CPSASE_8; 2.
 DR PROSITE; PS00617; CPSASE_9; 2.
 DR PROSITE; PS00618; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00619; CPSASE_1; 2.
 DR PROSITE; PS00620; CPSASE_2; 2.
 DR PROSITE; PS00621; CPSASE_3; 2.
 DR PROSITE; PS00622; CPSASE_4; 2.
 DR PROSITE; PS00623; CPSASE_5; 2.
 DR PROSITE; PS00624; CPSASE_6; 2.
 DR PROSITE; PS00625; CPSASE_7; 2.
 DR PROSITE; PS00626; CPSASE_8; 2.
 DR PROSITE; PS00627; CPSASE_9; 2.
 DR PROSITE; PS00628; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00629; CPSASE_1; 2.
 DR PROSITE; PS00630; CPSASE_2; 2.
 DR PROSITE; PS00631; CPSASE_3; 2.
 DR PROSITE; PS00632; CPSASE_4; 2.
 DR PROSITE; PS00633; CPSASE_5; 2.
 DR PROSITE; PS00634; CPSASE_6; 2.
 DR PROSITE; PS00635; CPSASE_7; 2.
 DR PROSITE; PS00636; CPSASE_8; 2.
 DR PROSITE; PS00637; CPSASE_9; 2.
 DR PROSITE; PS00638; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00639; CPSASE_1; 2.
 DR PROSITE; PS00640; CPSASE_2; 2.
 DR PROSITE; PS00641; CPSASE_3; 2.
 DR PROSITE; PS00642; CPSASE_4; 2.
 DR PROSITE; PS00643; CPSASE_5; 2.
 DR PROSITE; PS00644; CPSASE_6; 2.
 DR PROSITE; PS00645; CPSASE_7; 2.
 DR PROSITE; PS00646; CPSASE_8; 2.
 DR PROSITE; PS00647; CPSASE_9; 2.
 DR PROSITE; PS00648; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00649; CPSASE_1; 2.
 DR PROSITE; PS00650; CPSASE_2; 2.
 DR PROSITE; PS00651; CPSASE_3; 2.
 DR PROSITE; PS00652; CPSASE_4; 2.
 DR PROSITE; PS00653; CPSASE_5; 2.
 DR PROSITE; PS00654; CPSASE_6; 2.
 DR PROSITE; PS00655; CPSASE_7; 2.
 DR PROSITE; PS00656; CPSASE_8; 2.
 DR PROSITE; PS00657; CPSASE_9; 2.
 DR PROSITE; PS00658; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00659; CPSASE_1; 2.
 DR PROSITE; PS00660; CPSASE_2; 2.
 DR PROSITE; PS00661; CPSASE_3; 2.
 DR PROSITE; PS00662; CPSASE_4; 2.
 DR PROSITE; PS00663; CPSASE_5; 2.
 DR PROSITE; PS00664; CPSASE_6; 2.
 DR PROSITE; PS00665; CPSASE_7; 2.
 DR PROSITE; PS00666; CPSASE_8; 2.
 DR PROSITE; PS00667; CPSASE_9; 2.
 DR PROSITE; PS00668; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00669; CPSASE_1; 2.
 DR PROSITE; PS00670; CPSASE_2; 2.
 DR PROSITE; PS00671; CPSASE_3; 2.
 DR PROSITE; PS00672; CPSASE_4; 2.
 DR PROSITE; PS00673; CPSASE_5; 2.
 DR PROSITE; PS00674; CPSASE_6; 2.
 DR PROSITE; PS00675; CPSASE_7; 2.
 DR PROSITE; PS00676; CPSASE_8; 2.
 DR PROSITE; PS00677; CPSASE_9; 2.
 DR PROSITE; PS00678; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00679; CPSASE_1; 2.
 DR PROSITE; PS00680; CPSASE_2; 2.
 DR PROSITE; PS00681; CPSASE_3; 2.
 DR PROSITE; PS00682; CPSASE_4; 2.
 DR PROSITE; PS00683; CPSASE_5; 2.
 DR PROSITE; PS00684; CPSASE_6; 2.
 DR PROSITE; PS00685; CPSASE_7; 2.
 DR PROSITE; PS00686; CPSASE_8; 2.
 DR PROSITE; PS00687; CPSASE_9; 2.
 DR PROSITE; PS00688; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00689; CPSASE_1; 2.
 DR PROSITE; PS00690; CPSASE_2; 2.
 DR PROSITE; PS00691; CPSASE_3; 2.
 DR PROSITE; PS00692; CPSASE_4; 2.
 DR PROSITE; PS00693; CPSASE_5; 2.
 DR PROSITE; PS00694; CPSASE_6; 2.
 DR PROSITE; PS00695; CPSASE_7; 2.
 DR PROSITE; PS00696; CPSASE_8; 2.
 DR PROSITE; PS00697; CPSASE_9; 2.
 DR PROSITE; PS00698; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00699; CPSASE_1; 2.
 DR PROSITE; PS00700; CPSASE_2; 2.
 DR PROSITE; PS00701; CPSASE_3; 2.
 DR PROSITE; PS00702; CPSASE_4; 2.
 DR PROSITE; PS00703; CPSASE_5; 2.
 DR PROSITE; PS00704; CPSASE_6; 2.
 DR PROSITE; PS00705; CPSASE_7; 2.
 DR PROSITE; PS00706; CPSASE_8; 2.
 DR PROSITE; PS00707; CPSASE_9; 2.
 DR PROSITE; PS00708; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00709; CPSASE_1; 2.
 DR PROSITE; PS00710; CPSASE_2; 2.
 DR PROSITE; PS00711; CPSASE_3; 2.
 DR PROSITE; PS00712; CPSASE_4; 2.
 DR PROSITE; PS00713; CPSASE_5; 2.
 DR PROSITE; PS00714; CPSASE_6; 2.
 DR PROSITE; PS00715; CPSASE_7; 2.
 DR PROSITE; PS00716; CPSASE_8; 2.
 DR PROSITE; PS00717; CPSASE_9; 2.
 DR PROSITE; PS00718; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00719; CPSASE_1; 2.
 DR PROSITE; PS00720; CPSASE_2; 2.
 DR PROSITE; PS00721; CPSASE_3; 2.
 DR PROSITE; PS00722; CPSASE_4; 2.
 DR PROSITE; PS00723; CPSASE_5; 2.
 DR PROSITE; PS00724; CPSASE_6; 2.
 DR PROSITE; PS00725; CPSASE_7; 2.
 DR PROSITE; PS00726; CPSASE_8; 2.
 DR PROSITE; PS00727; CPSASE_9; 2.
 DR PROSITE; PS00728; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00729; CPSASE_1; 2.
 DR PROSITE; PS00730; CPSASE_2; 2.
 DR PROSITE; PS00731; CPSASE_3; 2.
 DR PROSITE; PS00732; CPSASE_4; 2.
 DR PROSITE; PS00733; CPSASE_5; 2.
 DR PROSITE; PS00734; CPSASE_6; 2.
 DR PROSITE; PS00735; CPSASE_7; 2.
 DR PROSITE; PS00736; CPSASE_8; 2.
 DR PROSITE; PS00737; CPSASE_9; 2.
 DR PROSITE; PS00738; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00739; CPSASE_1; 2.
 DR PROSITE; PS00740; CPSASE_2; 2.
 DR PROSITE; PS00741; CPSASE_3; 2.
 DR PROSITE; PS00742; CPSASE_4; 2.
 DR PROSITE; PS00743; CPSASE_5; 2.
 DR PROSITE; PS00744; CPSASE_6; 2.
 DR PROSITE; PS00745; CPSASE_7; 2.
 DR PROSITE; PS00746; CPSASE_8; 2.
 DR PROSITE; PS00747; CPSASE_9; 2.
 DR PROSITE; PS00748; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00749; CPSASE_1; 2.
 DR PROSITE; PS00750; CPSASE_2; 2.
 DR PROSITE; PS00751; CPSASE_3; 2.
 DR PROSITE; PS00752; CPSASE_4; 2.
 DR PROSITE; PS00753; CPSASE_5; 2.
 DR PROSITE; PS00754; CPSASE_6; 2.
 DR PROSITE; PS00755; CPSASE_7; 2.
 DR PROSITE; PS00756; CPSASE_8; 2.
 DR PROSITE; PS00757; CPSASE_9; 2.
 DR PROSITE; PS00758; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00759; CPSASE_1; 2.
 DR PROSITE; PS00760; CPSASE_2; 2.
 DR PROSITE; PS00761; CPSASE_3; 2.
 DR PROSITE; PS00762; CPSASE_4; 2.
 DR PROSITE; PS00763; CPSASE_5; 2.
 DR PROSITE; PS00764; CPSASE_6; 2.
 DR PROSITE; PS00765; CPSASE_7; 2.
 DR PROSITE; PS00766; CPSASE_8; 2.
 DR PROSITE; PS00767; CPSASE_9; 2.
 DR PROSITE; PS00768; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00769; CPSASE_1; 2.
 DR PROSITE; PS00770; CPSASE_2; 2.
 DR PROSITE; PS00771; CPSASE_3; 2.
 DR PROSITE; PS00772; CPSASE_4; 2.
 DR PROSITE; PS00773; CPSASE_5; 2.
 DR PROSITE; PS00774; CPSASE_6; 2.
 DR PROSITE; PS00775; CPSASE_7; 2.
 DR PROSITE; PS00776; CPSASE_8; 2.
 DR PROSITE; PS00777; CPSASE_9; 2.
 DR PROSITE; PS00778; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00779; CPSASE_1; 2.
 DR PROSITE; PS00780; CPSASE_2; 2.
 DR PROSITE; PS00781; CPSASE_3; 2.
 DR PROSITE; PS00782; CPSASE_4; 2.
 DR PROSITE; PS00783; CPSASE_5; 2.
 DR PROSITE; PS00784; CPSASE_6; 2.
 DR PROSITE; PS00785; CPSASE_7; 2.
 DR PROSITE; PS00786; CPSASE_8; 2.
 DR PROSITE; PS00787; CPSASE_9; 2.
 DR PROSITE; PS00788; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00789; CPSASE_1; 2.
 DR PROSITE; PS00790; CPSASE_2; 2.
 DR PROSITE; PS00791; CPSASE_3; 2.
 DR PROSITE; PS00792; CPSASE_4; 2.
 DR PROSITE; PS00793; CPSASE_5; 2.
 DR PROSITE; PS00794; CPSASE_6; 2.
 DR PROSITE; PS00795; CPSASE_7; 2.
 DR PROSITE; PS00796; CPSASE_8; 2.
 DR PROSITE; PS00797; CPSASE_9; 2.
 DR PROSITE; PS00798; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00799; CPSASE_1; 2.
 DR PROSITE; PS00800; CPSASE_2; 2.
 DR PROSITE; PS00801; CPSASE_3; 2.
 DR PROSITE; PS00802; CPSASE_4; 2.
 DR PROSITE; PS00803; CPSASE_5; 2.
 DR PROSITE; PS00804; CPSASE_6; 2.
 DR PROSITE; PS00805; CPSASE_7; 2.
 DR PROSITE; PS00806; CPSASE_8; 2.
 DR PROSITE; PS00807; CPSASE_9; 2.
 DR PROSITE; PS00808; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00809; CPSASE_1; 2.
 DR PROSITE; PS00810; CPSASE_2; 2.
 DR PROSITE; PS00811; CPSASE_3; 2.
 DR PROSITE; PS00812; CPSASE_4; 2.
 DR PROSITE; PS00813; CPSASE_5; 2.
 DR PROSITE; PS00814; CPSASE_6; 2.
 DR PROSITE; PS00815; CPSASE_7; 2.
 DR PROSITE; PS00816; CPSASE_8; 2.
 DR PROSITE; PS00817; CPSASE_9; 2.
 DR PROSITE; PS00818; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00819; CPSASE_1; 2.
 DR PROSITE; PS00820; CPSASE_2; 2.
 DR PROSITE; PS00821; CPSASE_3; 2.
 DR PROSITE; PS00822; CPSASE_4; 2.
 DR PROSITE; PS00823; CPSASE_5; 2.
 DR PROSITE; PS00824; CPSASE_6; 2.
 DR PROSITE; PS00825; CPSASE_7; 2.
 DR PROSITE; PS00826; CPSASE_8; 2.
 DR PROSITE; PS00827; CPSASE_9; 2.
 DR PROSITE; PS00828; CPSASE.
 DR TIGRFAMS; TIGR01367; CPSaseL_lrg; 1.
 DR PROSITE; PS00829; CPSASE_1; 2.
 DR PROSITE; PS00830; CPSASE_2; 2.
 DR PROSITE; PS00831; CPSASE_3; 2.
 DR PROSITE; PS00832; CPSASE_4; 2.
 DR PROSITE; PS00833; CPSASE_5; 2.
 DR PROSITE; PS00834;

DR PRINTS; PRO1541; GONADOLIBNI.
 DR PROSITE; PS00473; GNRH; 1.
 KW Amidation; Hormone; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 32 MGNRH ASSOCIATED PEPTIDE.
 FT CHAIN 33 91 BA15C9DC0843A7B CRC64;
 SQ SEQUENCE 91 AA; 9893 MW;

Query Match Score 52; DB 13; Length 91;
 Best Local Similarity 100.0%; Pred. No. 0.043%; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 HWSYGLRP 9
 Db 24 HWSYGLRP 31

RESULT 2
 Q8JIF3 PRELIMINARY; PRT; 64 AA.
 ID Q8JIF3 ; PRELIMINARY;
 AC Q8JIF3 ;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Gonadotropin-releasing hormone (Fragment).
 OS Dentex dentex.
 OC Actinopterygii; Neopterygii; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
 OC Sparidae; Dentex.
 OX NCBI_TaxID=94951;
 RN RP
 SEQUENCE FROM N.A.
 RA Kato M., Elmesiry G.E.;
 RT "Sequence comparison of GnRH genes in closely-related Sparidae
 fishes";
 RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB083313; BAC07229.1; -.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0005183; F:luteneizing hormone-releasing factor activity; IEA.
 DR InterPro; IPR002012; GnrH.
 DR InterPro; IPR004079; Gonadoliberini.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PRO1541; GONADOLIBNI.
 DR PROSITE; PS00473; GNRH; 1.
 FT NON TER 1 1
 FT NON TER 64 64 1028 MW;

Query Match Score 46; DB 13; Length 64;
 Best Local Similarity 87.5%; Pred. No. 0.37%; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 HWSYGLRP 9
 Db 16 HWSYGLSP 23

RESULT 3
 Q8JIF2 PRELIMINARY; PRT; 64 AA.
 ID Q8JIF2 ; PRELIMINARY;
 AC Q8JIF2 ;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Gonadotropin-releasing hormone (Fragment).
 OS Pagrus major (Red sea bream) (Chrysophryni major).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
 OC Sparidae; Acanthopagrus.
 OX NCBI_TaxID=81177;

RN RP
 SEQUENCE FROM N.A.
 RA Kato M., Elmesiry G.E.;
 RT "Sequence comparison of GnRH genes in closely-related Sparidae
 fishes";
 RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB083312; BAC07228.1; -.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0005183; F:luteneizing hormone-releasing factor activity; IEA.
 DR InterPro; IPR002012; GnrH.
 DR InterPro; IPR004079; Gonadoliberini.
 DR Pfam; PF00446; GnrH; 1.
 DR PRINTS; PRO1541; GONADOLIBNI.
 DR PROSITE; PS00473; GNRH; 1.
 FT NON TER 1 1
 FT NON TER 68 68 7543 MW;

Query Match Score 46; DB 13; Length 68;
 Best Local Similarity 87.5%; Pred. No. 0.4%; Indels 0; Gaps 0;

Qy 2 HWSYGLRP 9
 Db 20 HWSYGLSP 27

RESULT 5
 Q9YI26 PRELIMINARY;
 ID Q9YI26 PRELIMINARY;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW mode!

Run on: March 10, 2004, 08:58:54 ; Search time 9.88327 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-1
Perfect score: 54
Sequence: 1 XHWSYGLRPX 10

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:
 1: sp_arChaea: *
 2: sp_bacteria: *
 3: sp_fungi: *
 4: sp_human: *
 5: sp_invertebrate: *
 6: sp_mammal: *
 7: sp_mhc: *
 8: sp_organelle: *
 9: sp_phage: *
 10: sp_plant: *
 11: sp_rOdent: *
 12: sp_virus: *
 13: sp_vertebrate: *
 14: sp_unclassified: *
 15: sp_rvirus: *
 16: sp_bacteriAp: *
 17: sp_archeap: *

Result No.	Score	Query Match	Length	DB ID	Description
1	52	96.3	91	13 09PRHO	Oprh0 anguilla ja
2	46	85.2	64	13 Q8UIP3	Q8if3 dentex dent
3	46	85.2	64	13 Q8UIP2	Q8if2 pagrus major
4	46	85.2	68	13 Q8UIP4	Q8if4 acanthopag
5	46	85.2	87	13 Q9YI26	Q8if26 sparus aur
6	46	85.2	94	13 Q8JFY3	Q8if3 oreochromis
7	46	85.2	96	13 Q8UW80	Qauw80 verasper mo
8	46	85.2	98	13 Q805B0	Q805as oreochromis
9	46	85.2	120	13 Q805A5	Q805c micropogoni
10	43	79.6	374	16 Q9ZYR6	Q9yrs rhizostoma m
11	42	77.8	72	13 Q7T112	Q7T112 odontesthes
12	41	75.9	86	13 Q8JH60	Q8hb60 alosa sapid
13	41	75.9	1127	16 Q8GB15	Q88815 bifidobacte
14	40	74.1	164	5 Q86D87	Q8687 ciona intes
15	40	74.1	165	5 Q86D88	Q8688 ciona intes
16	40	74.1	367	16 Q987W1	Q987w1 rhizobium 1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	96.3	91	13 09PRHO	Oprh0 anguilla ja
2	46	85.2	64	13 Q8UIP3	Q8if3 dentex dent
3	46	85.2	64	13 Q8UIP2	Q8if2 pagrus major
4	46	85.2	68	13 Q8UIP4	Q8if4 acanthopag
5	46	85.2	87	13 Q9YI26	Q8if26 sparus aur
6	46	85.2	94	13 Q8JFY3	Q8if3 oreochromis
7	46	85.2	96	13 Q8UW80	Qauw80 verasper mo
8	46	85.2	98	13 Q805B0	Q805as oreochromis
9	46	85.2	120	13 Q805A5	Q805c micropogoni
10	43	79.6	374	16 Q9ZYR6	Q9yrs rhizostoma m
11	42	77.8	72	13 Q7T112	Q7T112 odontesthes
12	41	75.9	86	13 Q8JH60	Q8hb60 alosa sapid
13	41	75.9	1127	16 Q8GB15	Q88815 bifidobacte
14	40	74.1	164	5 Q86D87	Q8687 ciona intes
15	40	74.1	165	5 Q86D88	Q8688 ciona intes
16	40	74.1	367	16 Q987W1	Q987w1 rhizobium 1

ALIGNMENTS

RESULT 1	Q9PRHO	ID	Q9PRHO	PRELIMINARY	PRT;	91 AA.
	AC	Q9PRHO	ID	Q9PRHO	PRELIMINARY	
	DT	01-MAY-2000	DT	01-MAY-2000	(TREMBLE1.	13, Created)
					13, Last sequence update)	
					(TREMBLE1.	
					13, Last annotation update)	
					(TREMBLE1.	
					24, Last annotation update)	
					Prepro-mGnRH Precursor (Gonadotropin-releasing	
					DB hormone) (LH-RH) (Luteinizing Hormone)	
					OS Anguilla japonica (Japanese eel).	
					OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;	
					OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;	
					OC Anguilla.	
					OX NCBITaxID=7937;	
					RN [1]	
					RP SEQUENCE FROM N.A.	
					RC TISSUE=Brain , Suetake H , Aida K ;	
					RB Okubo K , Suetake H , Aida K ;	
					RT "A splicing variant for the prepro-mammalian gonadotropin-releasing	
					RT hormone (prepro-mGnRH) mRNA is present in the brain and various	
					RT peripheral tissues of the Japanese eel.";	
					RT genes in various tissues of the Japanese eel and evolution of GnRH. ";	
					RL Zool. Sci. 16:645-651(1999).	
					RT FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY	
					CC SIMILARITY: .	
					CC BELONGS TO THE GnRH FAMILY.	
					CC DR EMBL; AB026989; BAA226038; 1; -	
					CC DR EMBL; AB026991; BAA335971; -	
					CC DR GO; GO:0005576; C:extracellular; IEA.	
					CC DR GO; GO:0005183; P:uterinizing hormone-releasing factor activity; IEA.	
					CC DR GO; GO:0007275; P:development; IEA.	
					DR InterPro; IPR000012; GnRH.	
					DR InterPro; IPR004079; Gonadotropin.	
					DR Pfam; PF00446; GnRH.	

```

; FEATURE:
; OTHER INFORMATION: synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Glu is modified with a pyro group.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Gly is modified with -NH2 group.

US-10-184-126-1                               RESULT 15
Query Match                                     Query Match
Best Local Similarity                         Best Local Similarity
Matches   8; Conservative                      Matches   8; Conservative
          100.0%; Pred. No. 0.022;               100.0%; Pred. No. 0.022;
          Mismatches 0; Indels 0; Gaps 0;        Mismatches 0; Indels 0; Gaps 0;
Qy      2 HWSYGLRP 9                           Qy      2 HWSYGLRP 9
Db      2 HWSYGLRP 9                           Db      2 HWSYGLRP 9

US-10-117-364-1                               RESULT 15
Sequence 1, Application US/10117364
; Sequence 1, Application US/10117364
; Publication No. US20030181385A1
; GENERAL INFORMATION:
; APPLICANT: Roske, Roger W.
; TITLE OF INVENTION: LRRH Antagonist Peptides
; FILE REFERENCE: PPI-007CPUS
; CURRENT APPLICATION NUMBER: US/10/117,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/08/973,378
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/480,494
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-07
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-364-1                               Query Match
                                                Query Match
                                                Best Local Similarity 96.3%; Score 52; DB 14; Length 10;
                                                Matches 8; Conservative 100.0%; Pred. No. 0.022; Mismatches 0; Indels 0; Gaps 0;
Qy      2 HWSYGLRP 9                           Qy      2 HWSYGLRP 9
Db      2 HWSYGLRP 9                           Db      2 HWSYGLRP 9

Search completed: March 10, 2004, 10:25:47
Job time : 8.85592 secs

US-10-115-553-1                               RESULT 14
Query Match                                     Query Match
Best Local Similarity                         Best Local Similarity
Matches   8; Conservative                      Matches   8; Conservative
          96.3%; Score 52; DB 14; Length 10;
          100.0%; Pred. No. 0.022; Mismatches 0; Indels 0; Gaps 0;
Qy      2 HWSYGLRP 9                           Qy      2 HWSYGLRP 9
Db      2 HWSYGLRP 9                           Db      2 HWSYGLRP 9

US-10-122-483-1                               RESULT 14
Sequence 1, Application US/10122483
; Sequence 1, Application US/10122483
; Publication No. US20030044363A1
; GENERAL INFORMATION:
; APPLICANT: Ting, Chun-Jen
; APPLICANT: Hwang, Jau Lang
; APPLICANT: Hsu, Chia-Tse
; TITLE OF INVENTION: PEPTIDE REPEAT IMMUNOGENS
; FILE REFERENCE: 09919-071001
; CURRENT APPLICATION NUMBER: US/10/122,483
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: 09/412,558
; PRIOR FILING DATE: 1999-10-05
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-122-483-1

```

OX NCBI_TaxID=143350;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Okuzawa K., Grammenan J., Bogerd J., Goos H., Zohar Y., Kagawa H.;
 RL Submitted (SPP-1996) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D86582; BAA3129.1;
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; Gonadoliberini.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PRO1541; GONADOLIBRIN.
 DR PROSITE; PS00473; GnRH; 1.
 DR PROSITE; PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Signal; Multicene family; Pyroldone carboxylic acid.
 PT SIGNAL 1 23
 PT PEPTIDE 1 23
 PT CHAIN 24 95
 PT PEPTIDE 24 33
 PT PEPTIDE 37 95
 PT MOD_RES 24 24
 PT MOD_RES 33 33
 PT AMIDATION (G-34 PROVIDE AMIDE GROUP)
 PT (BY SIMILARITY).
 SQ SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;
 Query Match 85.2%; Score 46; DB 1; Length 95;
 Best Local Similarity 97.5%; Pred. No. 0.069;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 HNSYGLRP 9
 DB 25 HNSYGSP 32

Search completed: March 10, 2004, 09:13:45
 Job time : 2.82879 secs

[2]	SEQUENCE FROM N.A.	DT 28-FEB-2003 (Rel. 41, Last annotation update)
RP MEDLINE=99051842; PubMed=9143638;	DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GrRH-I).	
RX White R.B., Fernald R.D., "Ontogeny of Gonadotropin-releasing hormone (GrRH) gene expression reveals a distinct origin for GrRH-containing neurons in the midbrain," Gen. Comp. Endocrinol. 112:322-329 (1998).	DE (LH-RH I) (Lulliberin I).	
RN [3]	DT DE	
RP SEQUENCE OF 23-32, AND MASS SPECTROMETRY.	DN GNRH.	
RC TISSUE:pituitary;	OS Morone saxatilis (Striped bass); Craniata; Vertebrata; Euteleostomi; Bivalvia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Perciformes; Percoidae; Moronidae; Morone.	
RX MEDLINE=99051842; PubMed=7544702;	OC NCB-TaxID=34816;	
RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.B., White S.A., Francis R.C., Fernald R.D., Licht P., Warby C., Sherwood N.M.; "Primary structure of solitary form of gonadotropin-releasing hormone (GrRH) in cichlid pituitary; three forms of GrRH in brain of cichlid and pumpkinseed fish," Regul. Pept. 57:43-53 (1995).	OX RN	
RT -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALAMIC-PITUITARY-GONADAL AXIS.	[1] RP SEQUENCE FROM N.A. PubMed=9845669;	
CC -!- SUBCELLULAR LOCATION: Secreted.	RX MEDLINE=99061809; PubMed=9845669;	
CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOPHYSIAL AXONS.	RA Chow M.M., Knight K.E., Alok D., Stubblefield J., Zohar Y.; RT "Multiple GnRHs present in a teleost species are encoded by separate genes: analysis of the sGnRH and cGnRH-II genes from the striped bass, Morone saxatilis," J. Mol. Endocrinol. 21:277-289 (1998).	
CC -!- MASS SPECTROMETRY: MW=113.9; METHOD=VALDI; RANGE=23-32.	RT -!- FUNCTION: Stimulates the secretion of gonadotropins (By similarity).	
CC -!- SIMILARITY: Belongs to the GrRH family.	CC -!- SUBCELLULAR LOCATION: Secreted. Belongs to the GnRH family.	
CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	CC -!- SIMILARITY: Belongs to the GnRH family.	
CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC -!- DR EMBL; AF056311; AAC59691.1; -.	CC -!- DR EMBL; AF056311; AAC03817.1; -.	
CC -!- DR PIR; I50739; I50739.	CC -!- DR Intero; IPR002012; GrRH.	
CC -!- DR GO; GO:0005576; C:Extracellular; NAS.	CC -!- DR Pfam; PF00446; GrRH; 1.	
CC -!- DR GO; GO:0005183; F:Luteinizing hormone-releasing factor activity; NAS.	CC -!- DR PRINTS; PRO1541; CONADOLIBRIN1.	
CC -!- DR GO; GO:0007275; P:Development; IDA.	CC -!- DR PROSITE; PS00473; GrRH; 1.	
CC -!- DR InterPro; IPR002012; GrRH.	CC -!- DR KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Pyrolysis; Multigene family; Pyrolyzed carboxylic acid.	
CC -!- DR Pfam; PF00446; GrRH; 1.	CC -!- DR SIGNAL; 22 POTENTIAL.	
CC -!- DR PROSITE; PS00473; GrRH; 1.	CC -!- DR FT SIGNAL; 1 PROGNADOLIBERIN I.	
CC -!- DR KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Pyrolyzed carboxylic acid.	CC -!- DR FT CHAIN; 23 95 PROGNADOLIBERIN I.	
CC -!- DR SIGNAL; 1 22 PROGNADOLIBERIN I.	CC -!- DR FT PEPTIDE; 23 32 GNRH ASSOCIATED PEPTIDE I (POTENTIAL).	
CC -!- DR FT CHAIN; 23 94 PROGNADOLIBERIN I.	CC -!- DR FT PYRROLIDONE CARBOXYLIC ACID (BY SIMILARITY); 23 23 AMIDATION (G-33 PROVIDE AMIDE GROUP)	
CC -!- DR FT PEPTIDE; 23 94 GNRH ASSOCIATED PEPTIDE I (POTENTIAL).	CC -!- DR FT MOD_RES; 23 23 AMIDATION (G-33 PROVIDE AMIDE GROUP)	
CC -!- DR FT MOD_RES; 23 95 PYRROLIDONE CARBOXYLIC ACID.	CC -!- DR FT MOD_RES; 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP)	
CC -!- DR FT CONFFLICT; 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP).	CC -!- DR FT MOD_RES; 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP)	
CC -!- DR SQ SEQUENCE 94 AA; 10382 MW; E57DEB333278D7 CRC64;	CC -!- DR FT SIGNAL; 1 BY SIMILARITY; 98066988FC279FC CRC64;	
CC -!- DR QY 2 HWSYGLRP 9	CC -!- DR SQ SEQUENCE 95 AA; 10411 MW; 98066988FC279FC CRC64;	
CC -!- DR DB 24 HWSYGLSP 31	CC -!- DR QY 2 HWSYGLRP 9	
CC -!- DR DB 24 HWSYGLSP 31	CC -!- DR DB 24 HWSYGLSP 31	
CC -!- DR RESULT 15 GONI_PAGMA STANDARD; PRT; 95 AA.	CC -!- DR DB 24 HWSYGLSP 31	
CC -!- DR ID GONI_PAGMA	CC -!- DR ID GONI_PAGMA	
CC -!- DR AC P70074;	CC -!- DR AC P70074;	
CC -!- DT 15-JUL-1998 (Rel. 36, Created)	CC -!- DT 15-JUL-1998 (Rel. 36, Last sequence update)	
CC -!- DT 15-JUL-1998 (Rel. 41, Last annotation update)	CC -!- DT 28-FEB-2003 (Rel. 41, Last annotation update)	
CC -!- DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GrRH-I).	CC -!- DE Gonadoliberin I (Lulliberin I).	
CC -!- DE (LH-RH I) (Lulliberin I).	CC -!- DE (LH-RH I) (Lulliberin I).	
CC -!- GN GNRH.	CC -!- GN GNRH.	
CC -!- OS Parus major (Red sea bream) (Chrysophrys major).	CC -!- OS Parus major (Red sea bream) (Chrysophrys major).	
CC -!- OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Perciformes; Percoidae; Sparidae; Pagrus.	CC -!- OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Perciformes; Percoidae; Sparidae; Pagrus.	

DE Gonadotropin I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH I)
 DE (Iuliberin I).
 OS Alligator mississippiensis (American alligator).
 OC Archosauria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NCBI_TAXID=8496; Crocodylidae; Alligatorinae; Alligator.
 RN [1]
 RP
 RC TISSUE=Brain;
 RX MEDLINE=91352338; PubMed=1882082;
 RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
 RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure of two forms of gonadotropin-releasing hormone
 from brains of the American alligator (Alligator mississippiensis).";
 RL Regul Pept. 33:105-116(1991).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnrH family.
 DR PIR: A60006; REIAQ1;
 DR InterPro: IPR002012; GnRH.
 PFAM: PF00446; GNRH_1.
 PROSITE: PS00473; GNRH_1.
 DR Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.
 DR MOD_RES 1 PYRROLIDONE CARBOXYLIC ACID.
 FT AMIDATION
 SQ SEQUENCE 10 AA: 10 10 MW: 284B23D7256B45A3 CRC64;
 Query Match Score 48; DB 1; Length 10;
 Best Local Similarity 87.5%; Pred. No. 0.003;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRP 9
 Db 2 HWSYGLQP 9
 RESULT 12
 GONI_CHICK
 ID GONI_CHICK STANDARD; PRT; 92 AA.
 AC P37474; P20407;
 DT 01-JUN-1991 (Rel. 17, Created)
 DT 28-FEB-2003 (Rel. 29, Last sequence update)
 DE Progonadotropin I precursor [Contains: Gonadotropin I (LH-RH I)
 DE Luteinizing hormone-releasing hormone I] (Gondotropin-releasing
 hormone I) (GnRH I) (Iuliberin I); GnRH-associated Peptide I].
 DS Gallus gallus (Chicken).
 OC Bukaryota; Metazoa (Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 RN _TaxID=9031;
 RP SEQUENCE FROM N.A.
 RC STRAIN=White Leghorn;
 RX MEDLINE=94059455; PubMed=7992095;
 RA Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;
 RT "Characterization of the chicken preprogonadotrophin-releasing
 hormone I gene.";
 RL J. Mol. Endocrinol. 11:19-29(1993).
 RN [2]
 RP SEQUENCE OF 24-33.
 RC TISSUE=Hypothalamus;
 RX MEDLINE=822265778; PubMed=7050119;
 RA King J.A., Millar R.P.;
 RT "Structure of chicken hypothalamic luteinizing hormone-releasing
 hormone II. Isolation and characterization.";
 RL J. Biol. Chem. 257:10729-10732(1982).
 RN [3]
 RP SEQUENCE OF 24-33.
 RC TISSUE=Hypothalamus;
 RX MEDLINE=822265778; PubMed=7050119;
 RA King J.A., Millar R.P.;
 RT "Structure of avian hypothalamic gonadotrophin-releasing hormone.";
 RL S. Afr. J. Sci. 78:1124-1125(1982).

RN [4] SYNTHESIS OF 24-33.
 RP RX MEDLINE=822265777; PubMed=7050118;
 RA King J.A., Millar R.P.;
 RT "Structure of chicken hypothalamic luteinizing hormone-releasing
 hormone. I. Structural determination on partially purified
 material.";
 RL J. Biol. Chem. 257:10722-10728(1982).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnrH Family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC DR EMBL: X69491; CBA49246; 1.
 DR PIR: 150644; 150644.
 DR GO: GO:0005576; C: extracellular; IDA.
 DR GO: GO:0005183; P: luteinizing hormone-releasing factor activity; IDA.
 DR GO: GO:0007275; P: development; IDA.
 DR InterPro: IPR002012; GnrH.
 DR InterPro: IPR004079; Gonadotropin I.
 DR Pfam: PF00446; GnrH_1.
 DR PRINTS: PRO1541; GONADOLIBRN1.
 DR PROSITE: PS00474; GNRH_1.
 DR PROSITE: PS00475; GNRH_1.
 DR PROSITE: PS00477; GNRH_1.
 DR PROSITE: PS00478; GNRH_1.
 DR PROSITE: PS00479; GNRH_1.
 DR PROSITE: PS00480; GNRH_1.
 DR PROSITE: PS00481; GNRH_1.
 DR PROSITE: PS00482; GNRH_1.
 DR PROSITE: PS00483; GNRH_1.
 DR PROSITE: PS00484; GNRH_1.
 DR PROSITE: PS00485; GNRH_1.
 DR PROSITE: PS00486; GNRH_1.
 DR PROSITE: PS00487; GNRH_1.
 DR PROSITE: PS00488; GNRH_1.
 DR PROSITE: PS00489; GNRH_1.
 DR PROSITE: PS00490; GNRH_1.
 DR PROSITE: PS00491; GNRH_1.
 DR PROSITE: PS00492; GNRH_1.
 DR PROSITE: PS00493; GNRH_1.
 DR PROSITE: PS00494; GNRH_1.
 DR PROSITE: PS00495; GNRH_1.
 DR PROSITE: PS00496; GNRH_1.
 DR PROSITE: PS00497; GNRH_1.
 DR PROSITE: PS00498; GNRH_1.
 DR PROSITE: PS00499; GNRH_1.
 DR PROSITE: PS00500; GNRH_1.
 DR PROSITE: PS00501; GNRH_1.
 DR PROSITE: PS00502; GNRH_1.
 DR PROSITE: PS00503; GNRH_1.
 DR PROSITE: PS00504; GNRH_1.
 DR PROSITE: PS00505; GNRH_1.
 DR PROSITE: PS00506; GNRH_1.
 DR PROSITE: PS00507; GNRH_1.
 DR PROSITE: PS00508; GNRH_1.
 DR PROSITE: PS00509; GNRH_1.
 DR PROSITE: PS00510; GNRH_1.
 DR PROSITE: PS00511; GNRH_1.
 DR PROSITE: PS00512; GNRH_1.
 DR PROSITE: PS00513; GNRH_1.
 DR PROSITE: PS00514; GNRH_1.
 DR PROSITE: PS00515; GNRH_1.
 DR PROSITE: PS00516; GNRH_1.
 DR PROSITE: PS00517; GNRH_1.
 DR PROSITE: PS00518; P: luteinizing hormone-releasing factor activity; IDA.
 DR PROSITE: PS00519; GNRH_1.
 DR PROSITE: PS00520; GNRH_1.
 DR PROSITE: PS00521; GNRH_1.
 DR PROSITE: PS00522; GNRH_1.
 DR PROSITE: PS00523; GNRH_1.
 DR PROSITE: PS00524; GNRH_1.
 DR PROSITE: PS00525; GNRH_1.
 DR PROSITE: PS00526; GNRH_1.
 DR PROSITE: PS00527; GNRH_1.
 DR PROSITE: PS00528; GNRH_1.
 DR PROSITE: PS00529; GNRH_1.
 DR PROSITE: PS00530; GNRH_1.
 DR PROSITE: PS00531; GNRH_1.
 DR PROSITE: PS00532; GNRH_1.
 DR PROSITE: PS00533; GNRH_1.
 DR PROSITE: PS00534; GNRH_1.
 DR PROSITE: PS00535; GNRH_1.
 DR PROSITE: PS00536; GNRH_1.
 DR PROSITE: PS00537; GNRH_1.
 DR PROSITE: PS00538; GNRH_1.
 DR PROSITE: PS00539; GNRH_1.
 DR PROSITE: PS00540; GNRH_1.
 DR PROSITE: PS00541; GNRH_1.
 DR PROSITE: PS00542; GNRH_1.
 DR PROSITE: PS00543; GNRH_1.
 DR PROSITE: PS00544; GNRH_1.
 DR PROSITE: PS00545; GNRH_1.
 DR PROSITE: PS00546; GNRH_1.
 DR PROSITE: PS00547; GNRH_1.
 DR PROSITE: PS00548; GNRH_1.
 DR PROSITE: PS00549; GNRH_1.
 DR PROSITE: PS00550; GNRH_1.
 DR PROSITE: PS00551; GNRH_1.
 DR PROSITE: PS00552; GNRH_1.
 DR PROSITE: PS00553; GNRH_1.
 DR PROSITE: PS00554; GNRH_1.
 DR PROSITE: PS00555; GNRH_1.
 DR PROSITE: PS00556; GNRH_1.
 DR PROSITE: PS00557; GNRH_1.
 DR PROSITE: PS00558; GNRH_1.
 DR PROSITE: PS00559; GNRH_1.
 DR PROSITE: PS00560; GNRH_1.
 DR PROSITE: PS00561; GNRH_1.
 DR PROSITE: PS00562; GNRH_1.
 DR PROSITE: PS00563; GNRH_1.
 DR PROSITE: PS00564; GNRH_1.
 DR PROSITE: PS00565; GNRH_1.
 DR PROSITE: PS00566; GNRH_1.
 DR PROSITE: PS00567; GNRH_1.
 DR PROSITE: PS00568; GNRH_1.
 DR PROSITE: PS00569; GNRH_1.
 DR PROSITE: PS00570; GNRH_1.
 DR PROSITE: PS00571; GNRH_1.
 DR PROSITE: PS00572; GNRH_1.
 DR PROSITE: PS00573; GNRH_1.
 DR PROSITE: PS00574; GNRH_1.
 DR PROSITE: PS00575; GNRH_1.
 DR PROSITE: PS00576; GNRH_1.
 DR PROSITE: PS00577; GNRH_1.
 DR PROSITE: PS00578; GNRH_1.
 DR PROSITE: PS00579; GNRH_1.
 DR PROSITE: PS00580; GNRH_1.
 DR PROSITE: PS00581; GNRH_1.
 DR PROSITE: PS00582; GNRH_1.
 DR PROSITE: PS00583; GNRH_1.
 DR PROSITE: PS00584; GNRH_1.
 DR PROSITE: PS00585; GNRH_1.
 DR PROSITE: PS00586; GNRH_1.
 DR PROSITE: PS00587; GNRH_1.
 DR PROSITE: PS00588; GNRH_1.
 DR PROSITE: PS00589; GNRH_1.
 DR PROSITE: PS00590; GNRH_1.
 DR PROSITE: PS00591; GNRH_1.
 DR PROSITE: PS00592; GNRH_1.
 DR PROSITE: PS00593; GNRH_1.
 DR PROSITE: PS00594; GNRH_1.
 DR PROSITE: PS00595; GNRH_1.
 DR PROSITE: PS00596; GNRH_1.
 DR PROSITE: PS00597; GNRH_1.
 DR PROSITE: PS00598; GNRH_1.
 DR PROSITE: PS00599; GNRH_1.
 DR PROSITE: PS00600; GNRH_1.
 DR PROSITE: PS00601; GNRH_1.
 DR PROSITE: PS00602; GNRH_1.
 DR PROSITE: PS00603; GNRH_1.
 DR PROSITE: PS00604; GNRH_1.
 DR PROSITE: PS00605; GNRH_1.
 DR PROSITE: PS00606; GNRH_1.
 DR PROSITE: PS00607; GNRH_1.
 DR PROSITE: PS00608; GNRH_1.
 DR PROSITE: PS00609; GNRH_1.
 DR PROSITE: PS00610; GNRH_1.
 DR PROSITE: PS00611; GNRH_1.
 DR PROSITE: PS00612; GNRH_1.
 DR PROSITE: PS00613; GNRH_1.
 DR PROSITE: PS00614; GNRH_1.
 DR PROSITE: PS00615; GNRH_1.
 DR PROSITE: PS00616; GNRH_1.
 DR PROSITE: PS00617; GNRH_1.
 DR PROSITE: PS00618; GNRH_1.
 DR PROSITE: PS00619; GNRH_1.
 DR PROSITE: PS00620; GNRH_1.
 DR PROSITE: PS00621; GNRH_1.
 DR PROSITE: PS00622; GNRH_1.
 DR PROSITE: PS00623; GNRH_1.
 DR PROSITE: PS00624; GNRH_1.
 DR PROSITE: PS00625; GNRH_1.
 DR PROSITE: PS00626; GNRH_1.
 DR PROSITE: PS00627; GNRH_1.
 DR PROSITE: PS00628; GNRH_1.
 DR PROSITE: PS00629; GNRH_1.
 DR PROSITE: PS00630; GNRH_1.
 DR PROSITE: PS00631; GNRH_1.
 DR PROSITE: PS00632; GNRH_1.
 DR PROSITE: PS00633; GNRH_1.
 DR PROSITE: PS00634; GNRH_1.
 DR PROSITE: PS00635; GNRH_1.
 DR PROSITE: PS00636; GNRH_1.
 DR PROSITE: PS00637; GNRH_1.
 DR PROSITE: PS00638; GNRH_1.
 DR PROSITE: PS00639; GNRH_1.
 DR PROSITE: PS00640; GNRH_1.
 DR PROSITE: PS00641; GNRH_1.
 DR PROSITE: PS00642; GNRH_1.
 DR PROSITE: PS00643; GNRH_1.
 DR PROSITE: PS00644; GNRH_1.
 DR PROSITE: PS00645; GNRH_1.
 DR PROSITE: PS00646; GNRH_1.
 DR PROSITE: PS00647; GNRH_1.
 DR PROSITE: PS00648; GNRH_1.
 DR PROSITE: PS00649; GNRH_1.
 DR PROSITE: PS00650; GNRH_1.
 DR PROSITE: PS00651; GNRH_1.
 DR PROSITE: PS00652; GNRH_1.
 DR PROSITE: PS00653; GNRH_1.
 DR PROSITE: PS00654; GNRH_1.
 DR PROSITE: PS00655; GNRH_1.
 DR PROSITE: PS00656; GNRH_1.
 DR PROSITE: PS00657; GNRH_1.
 DR PROSITE: PS00658; GNRH_1.
 DR PROSITE: PS00659; GNRH_1.
 DR PROSITE: PS00660; GNRH_1.
 DR PROSITE: PS00661; GNRH_1.
 DR PROSITE: PS00662; GNRH_1.
 DR PROSITE: PS00663; GNRH_1.
 DR PROSITE: PS00664; GNRH_1.
 DR PROSITE: PS00665; GNRH_1.
 DR PROSITE: PS00666; GNRH_1.
 DR PROSITE: PS00667; GNRH_1.
 DR PROSITE: PS00668; GNRH_1.
 DR PROSITE: PS00669; GNRH_1.
 DR PROSITE: PS00670; GNRH_1.
 DR PROSITE: PS00671; GNRH_1.
 DR PROSITE: PS00672; GNRH_1.
 DR PROSITE: PS00673; GNRH_1.
 DR PROSITE: PS00674; GNRH_1.
 DR PROSITE: PS00675; GNRH_1.
 DR PROSITE: PS00676; GNRH_1.
 DR PROSITE: PS00677; GNRH_1.
 DR PROSITE: PS00678; GNRH_1.
 DR PROSITE: PS00679; GNRH_1.
 DR PROSITE: PS00680; GNRH_1.
 DR PROSITE: PS00681; GNRH_1.
 DR PROSITE: PS00682; GNRH_1.
 DR PROSITE: PS00683; GNRH_1.
 DR PROSITE: PS00684; GNRH_1.
 DR PROSITE: PS00685; GNRH_1.
 DR PROSITE: PS00686; GNRH_1.
 DR PROSITE: PS00687; GNRH_1.
 DR PROSITE: PS00688; GNRH_1.
 DR PROSITE: PS00689; GNRH_1.
 DR PROSITE: PS00690; GNRH_1.
 DR PROSITE: PS00691; GNRH_1.
 DR PROSITE: PS00692; GNRH_1.
 DR PROSITE: PS00693; GNRH_1.
 DR PROSITE: PS00694; GNRH_1.
 DR PROSITE: PS00695; GNRH_1.
 DR PROSITE: PS00696; GNRH_1.
 DR PROSITE: PS00697; GNRH_1.
 DR PROSITE: PS00698; GNRH_1.
 DR PROSITE: PS00699; GNRH_1.
 DR PROSITE: PS00700; GNRH_1.
 DR PROSITE: PS00701; GNRH_1.
 DR PROSITE: PS00702; GNRH_1.
 DR PROSITE: PS00703; GNRH_1.
 DR PROSITE: PS00704; GNRH_1.
 DR PROSITE: PS00705; GNRH_1.
 DR PROSITE: PS00706; GNRH_1.
 DR PROSITE: PS00707; GNRH_1.
 DR PROSITE: PS00708; GNRH_1.
 DR PROSITE: PS00709; GNRH_1.
 DR PROSITE: PS00710; GNRH_1.
 DR PROSITE: PS00711; GNRH_1.
 DR PROSITE: PS00712; GNRH_1.
 DR PROSITE: PS00713; GNRH_1.
 DR PROSITE: PS00714; GNRH_1.
 DR PROSITE: PS00715; GNRH_1.
 DR PROSITE: PS00716; GNRH_1.
 DR PROSITE: PS00717; GNRH_1.
 DR PROSITE: PS00718; GNRH_1.
 DR PROSITE: PS00719; GNRH_1.
 DR PROSITE: PS00720; GNRH_1.
 DR PROSITE: PS00721; GNRH_1.
 DR PROSITE: PS00722; GNRH_1.
 DR PROSITE: PS00723; GNRH_1.
 DR PROSITE: PS00724; GNRH_1.
 DR PROSITE: PS00725; GNRH_1.
 DR PROSITE: PS00726; GNRH_1.
 DR PROSITE: PS00727; GNRH_1.
 DR PROSITE: PS00728; GNRH_1.
 DR PROSITE: PS00729; GNRH_1.
 DR PROSITE: PS00730; GNRH_1.
 DR PROSITE: PS00731; GNRH_1.
 DR PROSITE: PS00732; GNRH_1.
 DR PROSITE: PS00733; GNRH_1.
 DR PROSITE: PS00734; GNRH_1.
 DR PROSITE: PS00735; GNRH_1.
 DR PROSITE: PS00736; GNRH_1.
 DR PROSITE: PS00737; GNRH_1.
 DR PROSITE: PS00738; GNRH_1.
 DR PROSITE: PS00739; GNRH_1.
 DR PROSITE: PS00740; GNRH_1.
 DR PROSITE: PS00741; GNRH_1.
 DR PROSITE: PS00742; GNRH_1.
 DR PROSITE: PS00743; GNRH_1.
 DR PROSITE: PS00744; GNRH_1.
 DR PROSITE: PS00745; GNRH_1.
 DR PROSITE: PS00746; GNRH_1.
 DR PROSITE: PS00747; GNRH_1.
 DR PROSITE: PS00748; GNRH_1.
 DR PROSITE: PS00749; GNRH_1.
 DR PROSITE: PS00750; GNRH_1.
 DR PROSITE: PS00751; GNRH_1.
 DR PROSITE: PS00752; GNRH_1.
 DR PROSITE: PS00753; GNRH_1.
 DR PROSITE: PS00754; GNRH_1.
 DR PROSITE: PS00755; GNRH_1.
 DR PROSITE: PS00756; GNRH_1.
 DR PROSITE: PS00757; GNRH_1.
 DR PROSITE: PS00758; GNRH_1.
 DR PROSITE: PS00759; GNRH_1.
 DR PROSITE: PS00760; GNRH_1.
 DR PROSITE: PS00761; GNRH_1.
 DR PROSITE: PS00762; GNRH_1.
 DR PROSITE: PS00763; GNRH_1.
 DR PROSITE: PS00764; GNRH_1.
 DR PROSITE: PS00765; GNRH_1.
 DR PROSITE: PS00766; GNRH_1.
 DR PROSITE: PS00767; GNRH_1.
 DR PROSITE: PS00768; GNRH_1.
 DR PROSITE: PS00769; GNRH_1.
 DR PROSITE: PS00770; GNRH_1.
 DR PROSITE: PS00771; GNRH_1.
 DR PROSITE: PS00772; GNRH_1.
 DR PROSITE: PS00773; GNRH_1.
 DR PROSITE: PS00774; GNRH_1.
 DR PROSITE: PS00775; GNRH_1.
 DR PROSITE: PS00776; GNRH_1.
 DR PROSITE: PS00777; GNRH_1.
 DR PROSITE: PS00778; GNRH_1.
 DR PROSITE: PS00779; GNRH_1.
 DR PROSITE: PS00780; GNRH_1.
 DR PROSITE: PS00781; GNRH_1.
 DR PROSITE: PS00782; GNRH_1.
 DR PROSITE: PS00783; GNRH_1.
 DR PROSITE: PS00784; GNRH_1.
 DR PROSITE: PS00785; GNRH_1.
 DR PROSITE: PS00786; GNRH_1.
 DR PROSITE: PS00787; GNRH_1.
 DR PROSITE: PS00788; GNRH_1.
 DR PROSITE: PS00789; GNRH_1.
 DR PROSITE: PS00790; GNRH_1.
 DR PROSITE: PS00791; GNRH_1.
 DR PROSITE: PS00792; GNRH_1.
 DR PROSITE: PS00793; GNRH_1.
 DR PROSITE: PS00794; GNRH_1.
 DR PROSITE: PS00795; GNRH_1.
 DR PROSITE: PS00796; GNRH_1.
 DR PROSITE: PS00797; GNRH_1.
 DR PROSITE: PS00798; GNRH_1.
 DR PROSITE: PS00799; GNRH_1.
 DR PROSITE: PS00800; GNRH_1.
 DR PROSITE: PS00801; GNRH_1.
 DR PROSITE: PS00802; GNRH_1.
 DR PROSITE: PS00803; GNRH_1.
 DR PROSITE: PS00804; GNRH_1.
 DR PROSITE: PS00805; GNRH_1.
 DR PROSITE: PS00806; GNRH_1.
 DR PROSITE: PS00807; GNRH_1.
 DR PROSITE: PS00808; GNRH_1.
 DR PROSITE: PS00809; GNRH_1.
 DR PROSITE: PS00810; GNRH_1.
 DR PROSITE: PS00811; GNRH_1.
 DR PROSITE: PS00812; GNRH_1.
 DR PROSITE: PS00813; GNRH_1.
 DR PROSITE: PS00814; GNRH_1.
 DR PROSITE: PS00815; GNRH_1.
 DR PROSITE: PS00816; GNRH_1.
 DR PROSITE: PS00817; GNRH_1.
 DR PROSITE: PS00818; GNRH_1.
 DR PROSITE: PS00819; GNRH_1.
 DR PROSITE: PS00820; GNRH_1.
 DR PROSITE: PS00821; GNRH_1.
 DR PROSITE: PS00822; GNRH_1.
 DR PROSITE: PS00823; GNRH_1.
 DR PROSITE: PS00824; GNRH_1.
 DR PROSITE: PS00825; GNRH_1.
 DR PROSITE: PS00826; GNRH_1.
 DR PROSITE: PS00827; GNRH_1.
 DR PROSITE: PS00828; GNRH_1.
 DR PROSITE: PS00829; GNRH_1.
 DR PROSITE: PS00830; GNRH_1.
 DR PROSITE: PS00831; GNRH_1.
 DR PROSITE: PS00832; GNRH_1.
 DR PROSITE: PS00833; GNRH_1.
 DR PROSITE: PS00834; GNRH_1.
 DR PROSITE: PS00835; GNRH_1.
 DR PROSITE: PS00836; GNRH_1.
 DR PROSITE: PS00837; GNRH_1.
 DR PROSITE: PS00838; GNRH_1.
 DR PROSITE: PS00839; GNRH_1.
 DR PROSITE: PS00840; GNRH_1.
 DR PROSITE: PS00841; GNRH_1.
 DR PROSITE: PS00842; GNRH_1.
 DR PROSITE: PS00843; GNRH_1.
 DR PROSITE: PS00844; GNRH_1.
 DR PROSITE: PS00845; GNRH_1.
 DR PROSITE: PS00846; GNRH_1.
 DR PROSITE: PS00847; GNRH_1.
 DR PROSITE: PS00848; GNRH_1.
 DR PROSITE: PS00849; GNRH_1.
 DR PROSITE: PS00850; GNRH_1.
 DR PROSITE: PS00851; GNRH_1.
 DR PROSITE: PS00852; GNRH_1.
 DR PROSITE: PS00853; GNRH_1.
 DR PROSITE: PS00854; GNRH_1.
 DR PROSITE: PS00855; GNRH_1.
 DR PROSITE: PS00856; GNRH_1.
 DR PROSITE: PS00857; GNRH_1.
 DR PROSITE: PS00858; GNRH_1.
 DR PROSITE: PS00859; GNRH_1.
 DR PROSITE: PS00860; GNRH_1.
 DR PROSITE: PS00861; GNRH_1.
 DR PROSITE: PS00862; GNRH_1.
 DR PROSITE: PS00863; GNRH_1.
 DR PROSITE: PS00864; GNRH_1.
 DR PROSITE: PS00865; GNRH_1.
 DR PROSITE: PS00866; GNRH_1.
 DR PROSITE: PS00867; GNRH_1.
 DR PROSITE: PS00868; GNRH_1.
 DR PROSITE: PS00869; GNRH_1.
 DR PROSITE: PS00870; GNRH_1.
 DR PROSITE: PS00871; GNRH_1.
 DR PROSITE: PS00872; GNRH_1.
 DR PROSITE: PS00873; GNRH_1.
 DR PROSITE: PS00874; GNRH_1.
 DR PROSITE: PS00875; GNRH_1.
 DR PROSITE: PS00876; GNRH_1.
 DR PROSITE: PS00877; GNRH_1.
 DR PROSITE: PS00878; GNRH_1.
 DR PROSITE: PS00879; GNRH_1.
 DR PROSITE: PS00880; GNRH_1.
 DR PROSITE: PS00881; GNRH_1.
 DR PROSITE: PS00882; GNRH_1.
 DR PROSITE: PS00883; GNRH_1.
 DR PROSITE: PS00884; GNRH_1.
 DR PROSITE: PS00885; GNRH_1.
 DR PROSITE: PS00886; GNRH_1.
 DR PROSITE: PS00887; GNRH_1.
 DR PROSITE: PS00888; GNRH_1.
 DR PROSITE: PS00889; GNRH_1.
 DR PROSITE: PS00890; GNRH_1.
 DR PROSITE: PS00891; GNRH_1.
 DR PROSITE: PS00892; GNRH_1.
 DR PROSITE: PS00893; GNRH_1.
 DR PROSITE: PS00894; GNRH_1.
 DR PROSITE: PS00895; GNRH_1.
 DR PROSITE: PS00896; GNRH_1.
 DR PROSITE: PS00897; GNRH_1.
 DR PROSITE: PS00898; GNRH_1.
 DR PROSITE: PS00899; GNRH_1.
 DR PROSITE: PS00900; GNRH_1.
 DR PROSITE: PS00901; GNRH_1.
 DR PROSITE: PS00902; GNRH_1.
 DR PROSITE: PS00903; GNRH_1.
 DR PROSITE: PS00904; GNRH_1.
 DR PROSITE: PS00905; GNRH_1.
 DR PROSITE: PS00906; GNRH_1.
 DR PROSITE: PS00907; GNRH_1.
 DR PROSITE: PS00908; GNRH_1.
 DR PROSITE: PS00909; GNRH_1.
 DR PROSITE: PS00910; GNRH_1.
 DR PROSITE: PS00911; GNRH_1.
 DR PROSITE: PS00912; GNRH_1.
 DR PROSITE: PS00913; GNRH_1.
 DR PROSITE: PS00914; GNRH_1.
 DR PROSITE: PS00915; GNRH_1.
 DR PROSITE: PS00916; GNRH_1.
 DR PROSITE: PS00917; GNRH_1.
 DR PROSITE: PS00918; GNRH_1.
 DR PROSITE: PS00919; GNRH_1.
 DR PROSITE: PS00920; GNRH_1.
 DR PROSITE: PS00921; GNRH_1.
 DR PROSITE: PS00922; GNRH_1.
 DR PROSITE: PS00923; GNRH_1.
 DR PROSITE: PS00924; GNRH_1.
 DR PROSITE: PS00925; GNRH_1.
 DR PROSITE: PS00926; GNRH_1.
 DR PROSITE: PS00927; GNRH_1.
 DR PROSITE: PS00928; GNRH_1.
 DR PROSITE: PS00929; GNRH_1.
 DR PROSITE: PS00930; GNRH_1.
 DR PROSITE: PS00931; GNRH_1.
 DR PROSITE: PS00932; GNRH_1.
 DR PROSITE: PS00933; GNRH_1.
 DR PROSITE: PS00934; GNRH_1.
 DR PROSITE: PS00935; GNRH_1.
 DR PROSITE: PS00936; GNRH_1.
 DR PROSITE: PS00937; GNRH_1.
 DR PROSITE: PS00938; GNRH_1.
 DR PROSITE: PS00939; GNRH_1.
 DR PROSITE: PS00940; GNRH_1.
 DR PROSITE: PS00941; GNRH_1.
 DR PROSITE: PS00942; GNRH_1.
 DR PROSITE: PS00943; GNRH_1.
 DR PROSITE: PS00944; GNRH_1.
 DR PROSITE: PS

[2] SEQUENCE FROM N.A.
RN RP STANDARD; PRT; 92 AA.
RN MEDLINE=8934461; PubMed=1476669;
RX Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.; structure and
RT "the first gonadotropin-releasing hormone: SH locus: structure and
hypothalamic expression." /
RL Mol. Endocrinol. 3:1257-1262 (1989).
RN [3]
RN SEQUENCE FROM N.A.
RN TISSUE=Thymus;
RX MEDLINE=93105480; PubMed=1468115;
RA Maier C.C., Marchetti B., Lebeau R.D., Blalock J.E.; structure and
RT "Thymocytes express a mRNA that is identical to hypothalamic
luteinizing hormone-releasing hormone mRNA." /
RL Cell. Mol. Neurobiol. 12:147-154 (1992).
RN [4]
RN SEQUENCE OF 1-47 FROM N.A.
RN TISSUE=Heart;
RX MEDLINE=87149087; PubMed=1547652;
PA Adelman J.P.; Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
DNA locus." /
RT Science 235:1514-1517(1987).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
the secretion of both luteinizing and follie-stimulating
hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Central nervous system.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC ---
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC ---
EMBL; S50870; AB24572; 1;
EMBL; M12579; AAX41263; 1;
DR InterPro; IPR004019; Gonadoliberin.
DR EMBL; M31670; AAX41261; 1;
DR EMBL; M11670; AAX41261; 1; ALT_SEQ.
DR EMBL; M15527; AAX42141; 1;
DR EMBL; M15529; AAX42139; 1;
DR PROSITE; PS00473; GNRH; 1.
DR PIR; A00147; RRHNG.
DR InterPro; IPR002012; Gnrh.
DR InterPro; IPR004019; Gonadoliberin.
DR Pfam; PF00446; Gnrh; 1.
DR PRINTS; PRO1541; GONADOLIBRNI.
DR PROSITE; PS00473; GNRH; 1.
DR ACT_SITE; 26
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
Placenta; Signal; Pyrrolidone carboxylic acid.
STINAL_ 1 23 PROGONDOLIBERIN I.
FT CHAIN 24 92 GONADOLIBERIN I.
FT PEPTIDE 24 33 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT PEPTIDE 37 92 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY.
FT ACT_SITE 26 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 24 24 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 92 AA; 10500 MW; 494B5664DA8A3EB3 CRC64;
Query Match 96.3%; Score 52; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0 005; Indels 0; Gaps 0;
Matches 8; Conservative 0; Missmatches 0; Indels 0; Gaps 0;
Qy 2 HWSYGLRP 9
Db 25 HWSYGLRP 32
RESULT 10
GON_TUPGB

[2] SEQUENCE FROM N.A.
RN RP STANDARD; PRT; 92 AA.
RN MEDLINE=97079639; PubMed=8921350;
RX Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGnRH mRNAs in the tree shrew:
first direct evidence for mesencephalic GnRH gene expression in a
placental mammal." /
RT Gen. Comp. Endocrinol. 104:7-19(1996).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
the secretion of both luteinizing and follie-stimulating
hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC ---
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC ---
DR EMBL; U63326; AAB6837; 1;
DR InterPro; IPR002012; Gnrh.
DR InterPro; IPR004019; Gonadoliberin.
DR Pfam; PF00446; Gnrh; 1.
DR PRINTS; PRO1541; GONADOLIBRNI.
DR PROSITE; PS00473; GNRH; 1.
DR ACT_SITE 26 26 Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 24 92 PROGONDOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
SIMILARITY).
SQ SEQUENCE 92 AA; 10197 MW;
Query Match 95.3%; Score 52; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0 005;
Matches 8; Conservative 0; Missmatches 0; Indels 0; Gaps 0;
Qy 2 HWSYGLRP 9
Db 25 HWSYGLRP 32
RESULT 11
GON1_ALIMI
ID GON1_ALIMI STANDARD; PRT; 10 AA.
AC P37041; P00407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

Page 5

DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Progonadotropin I precursor [Contains: Gonadotropin I (LHRH I)
 DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
 DE hormone I) (GnRH I); GnRH-associated peptide I (GAP1)].
 DB GNRH OR GNRH.
 GN Rana catesbeiana (Bull frog); Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Bokaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
 OC NCBTAXID=8400;
 RN [1] _SEQUENCE FROM N.A.; TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RP MEDLINE=21102951; PubMed=11170016;
 RC TISSUE=forebrain; RA Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
 RX Kwon H.B.; RT "Cloning and characterization of cDNAs encoding the GnRH1 and GnRH2
 precursors from bullfrog (*Rana catesbeiana*).";
 RT J. Exp. Zool. 289:190-201 (2001).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Forebrain.
 CC -!- DEVELOPMENTAL STAGE: Expressed at significantly higher levels
 CC during post-breeding. Not expressed in pituitary.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; API:18754; AAL05972.1; DR GO; GO:0005576; C:extracellular; NAS.
 CC GO; GO:0005183; F:luteinizing hormone-releasing factor activity; NAS.
 CC DR GO; GO:0009755; P:hormone mediated signaling; NAS.
 CC DR GO; GO:0000003; P:reproduction; NAS.
 CC DR InterPro; IPR002012; GnRH.
 CC DR Pfam; PF00446; GnRH_1.
 CC DR PRINTS; PRO1541; GONADOLIBRNI.
 CC DR PROSITE; PS00473; GnRH_1.
 CC DR Cleavage on pair of basic residues; Hormone; Amidation; Signal;
 KW Pyrrolidone carboxylic acid.
 FT SIGNAL 1 24 POTENTIAL, PROGONADOLIBERIN I.
 FT CHAIN 25 90 GONADOLIBERIN I.
 FT PEPTIDE 25 34 GNRH-ASSOCIATED PEPTIDE I (BY
 FT PEPTIDE 38 86 SIMILARITY).
 FT MOD_RES 25 25 PYRROLIDONE CARBOXYLIC ACID (BY
 FT MOD_RES 34 34 SIMILARITY).
 FT SEQUENCE 90 AA; 10291 MW; 317203B4B3DA2FEB7 CRC64;

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.0049; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DP 26 HWSYGLRP 33

RESULT 7
 GONI_PIG STANDARD; PRT; 91 AA.
 AC P49921;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

Query Match	96.3%	Score 52;	DB 1;	Length 67;				
Best Local Similarity	100.0%	Pred. No.	0.0036;	Indels	0;	Gaps	0;	
Matches	8	Conservative	0;	Mismatches	0;			
Qy	2 HWSYGLRP 9							
Db	7 HWSYGLRP 14							
RESULT 4								
GONI_XENLA	STANDARD;	PRT;	89 AA.					
AC P41656;								
ID GONI_XENLA								
OS Xenopus laevis (African clawed frog);								
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;								
CC Xanopoda; Xenopus.								
CN [1]								
RN RP								
TISSUE: Forebrain; PubMed=8137750;								
RA Hayes W.P.; Wray S.; Battye J.F.;								
RR "The frog gonadotropin-releasing hormone-I (GnRH-I) gene has a mammalian-like expression pattern and conserved domains in GnRH-associated peptide, but brain onset is delayed until metamorphosis".								
RT Endocrinology 134:1835-1844(1994).								
CC -!- FUNCTION: Stimulates the secretion of gonadotropins.								
CC -!- SUBCELLULAR LOCATION: Secreted.								
CC -!- SIMILARITY: Belongs to the GnRH family.								
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).								
CC EMBL; I2B040; PNA49728.1;								
DR PIR; I51123; I51423.								
DR InterPro; IPR002012; GnRH.								
DR InterPro; IPR002012; Gonadoliberin I.								
DR PFAM; PF00446; GnRH; 1.								
DR PRINTS; PS00473; GONADOLIBRNT.								
DR CLEAVAGE ON PAIR OF BASIC RESIDUES; Hormone; Amidation; Hypothalamus; SIGNAL; Pyrrolidone carboxylic acid.								
KW SIGNAL	1							
FT CHAIN	22							
FT PEPTIDE	22							
FT PEPTIDE	31							
FT ACT_SITE	24							
FT MOD_RES	22							
FT MOD_RES	31							
FT SEQUENCE	90 AA;	10337 MW;						
DR PROSIGNADOLIBERIN I.								
DR GONADOLIBERIN I.								
DR PROLACTIN RELEASE-INHIBITING FACTOR I.								
DR ACTIVITY.								
DR PYRROLIDONE CARBOXYLIC ACID.								
DR AMIDES (G-32 PROVIDE AMIDE GROUP).								
DR 1C0766FA48226ED9 CRC64;								
DR Score 52; DB 1; Length 90;								
DR Pred. No. 0.0049;								
DR Mismatches 0;								
DR Indels 0;								
DR Gaps 0;								
RESULT 5								
GONI_MOUSE	STANDARD;	PRT;	90 AA.					
ID GONI_MOUSE								
AC P13562;								
DT 01-JAN-1990 (Rel. 13, Last sequence update)								
DT 01-JAN-2003 (Rel. 41, Last annotation update)								
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)]								
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing factor hormone I) (GnRH I) (Lutiberin I); Prolactin release-inhibiting factor I].								
DE GNRH1 OR GNRH.								
GN Mus musculus (Mouse).								
OS Mus musculus (Mouse).								
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Murinae; Mus.								
OC NCBII_TaxID:10090;								
OX [1].								
RN								
RP SEQUENCE FROM N.A.								
RX MEDLINE=7065928; PubMed=3024317;								
RA Mason A.J.; Hayflick J.S.; Zoeller R.T.; Young W.S. III,								
RA Phillips H.S.; Nikolic K.; Seuberg P.H.;								
RT "A deletion truncating the gonadotropin-releasing hormone gene is responsible for hypogonadism in the hpg mouse.";								
RL Science 234:1366-1371(1986).								
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates the secretion of both luteinizing and follicle-stimulating hormones.								
CC -!- SUBCELLULAR LOCATION: Secreted.								
CC -!- SIMILARITY: Belongs to the GnRH Family.								
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).								
CC EMBL; ML4372; AAA37171.1; -.								
DR PIR; A47578; RHMSG.								
DR MGI; MGI_95789; GnRH.								
DR InterPro; IPR002012; GnRH.								
DR InterPro; IPR004079; Gonadoliberin I.								
DR PFAM; PF00446; GnRH; 1.								
DR PRINTS; PS00473; GONADOLIBRNT.								
DR PROSIGNADOLIBERIN I.								
DR GONADOLIBERIN I.								
DR PROLACTIN RELEASE-INHIBITING FACTOR I.								
DR ACTIVITY.								
DR PYRROLIDONE CARBOXYLIC ACID.								
DR AMIDES (G-32 PROVIDE AMIDE GROUP).								
DR 1C0766FA48226ED9 CRC64;								
DR Score 52; DB 1; Length 90;								
DR Pred. No. 0.0049;								
DR Mismatches 0;								
DR Indels 0;								
DR Gaps 0;								
RESULT 6								
GONI_RANCA	STANDARD;	PRT;	90 AA.					
ID GONI_RANCA								
AC Q9Y63;								
DT 10-OCT-2003 (Rel. 42, Created)								
DT 10-OCT-2003 (Rel. 42, Last sequence update)								
Query Match	96.3%	Score 52;	DB 1;	Length 89;				
Best Local Similarity	100.0%	Pred. No.	0.0048;	Indels	0;	Gaps	0;	
Matches	8	Conservative	0;	Mismatches	0;			
Qy	2 HWSYGLRP 9							
Db	25 HWSYGLRP 32							

PT	PEPTIDE	14	>61	GNRH-ASSOCIATED PEPTIDE I.	Score 52;	DB 1;	Length 63;
PT	ACT_SITE	3	3	APPEARS TO BE ESSENTIAL FOR BIOLOGICAL	Best Local Similarity	100.0%;	Pred. No. 0.0014;
PT	MOD_RES	1	1	ACTIVITY.	Matches	0;	Mismatches 0;
PT	MOD_RES	10	10	PRORROLIDONE CARBOXYLIC ACID.	Indels	0;	Gaps 0;
PT	NON_TER	61	61	AMIDATION (G-11 PROVIDE AMIDE GROUP).			
SQ	SEQUENCE	61 AA:	6828 MN;	63962AAE319B8F0 CRC64;			
Query Match Score	96.3%;	DB 1;	Length 61;				
Best Local Similarity	100.0%;	Pred. No. 0.0033;	Mismatches 0;				
Matches	8;	Conservative 0;	Indels 0;				
QY		2 HWSYGLRP 9		RESULT 3	GONI_MACMU	STANDARD:	PRT;
DB		2 HWSYGLRP 9		ID_GONI_MACMU			67 AA.
RESULT 2	GONI_MESAU	STANDARD:	PRT;	AC	P5347;		
AC	009163;			DT	01-OCT-1996 (Rel. 34, Created)		
DT	15-DEC-1998 (Rel. 37, Created)			DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)			DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Progonadotiberin I precursor [Contains: Gonadotiberin I (LH-RH I)			DE	Progonadotiberin I precursor [Contains: Gonadotiberin I (LH-RH I)		
DE	(Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing			DE	(Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing		
DE	hormone I) (GnRH I) (Lutiberin I); GnRH-associated Peptide I]			DE	hormone I) (GnRH I) (Lutiberin I); GnRH-associated Peptide I]		
DE	(Fragment).			DB	(Fragment).		
GN	GNRH OR GNRH OR LHRH			GN	GNRH OR LHRH OR LHRH		
GS	Mesocricetus auratus (Golden hamster)			GS	Mesocricetus auratus (Golden hamster)		
OC	Vertebrata; Chordata; Craniata; Euteleostomi;			OC	Macaca mulatta (Rhesus macaque)		
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;			OC	Bukaryota; Metazoa; Chordata; Craniata; Euteleostomi;		
OC	Cercopitheciinae; Macaca.			OC	Macaca mulatta (Rhesus macaque)		
NCBI_TaxID	9544;			NCBI_TaxID	9544;		
RN	[1]			RN	[1]		
RP	SEQUENCE FROM N.A.			RP	SEQUENCE FROM N.A.		
RA	Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;			RC	TISSUE=Hypothalamus;		
RL	Submitted (MAR1997) to the EMBL/GenBank/DBJ databases.			RX	Medline=95124501; PubMed=7545971;		
CC	-!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates			RA	Ma Y.J., Costa M.B., Oleda S.R.;		
CC	the secretion of both luteinizing and follicle-stimulating			RA	Developmental expression of the genes encoding transforming growth		
CC	hormones.			RT	factor alpha and its receptor in the hypothalamus of female rhesus		
CC	-!- SUBCELLULAR LOCATION: Secreted.			RT	macaques.;		
CC	-!- SIMILARITY: Belongs to the GnRH family.			RT	Neuroendocrinology 60:346-350(1994).		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			RL	-!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			CC	the secretion of both luteinizing and follicle-stimulating		
CC	the European Bioinformatics Institute. There are no restrictions on its			CC	hormones.		
CC	use by non-profit institutions as long as its content is in no way			CC	-!- SIMILARITY: Belongs to the GnRH family.		
CC	modified and this statement is not removed. Usage by and for commercial			CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	or send an email to license@isb-sib.ch).			CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way			CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial			CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).			CC	or send an email to license@isb-sib.ch).		
CC	EMBL; U91938; AB51302.1. -			CC	or send an email to license@isb-sib.ch).		
DR	InterPro; IPR000102; GnrH.			CC	or send an email to license@isb-sib.ch).		
DR	InterPro; IPR0001079; Gonadotiberin.			CC	or send an email to license@isb-sib.ch).		
DR	PFAM; PF00446; GnrH.			CC	or send an email to license@isb-sib.ch).		
DR	PRINTS; PRO1541; GONADOLIBERNI.			CC	or send an email to license@isb-sib.ch).		
DR	PROSITE; PS00473; GNRH_1.			CC	or send an email to license@isb-sib.ch).		
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;			CC	or send an email to license@isb-sib.ch).		
KW	Placenta; Pyrrolidone carboxylic acid.			CC	or send an email to license@isb-sib.ch).		
FT	NON_TER	1	1	FT	NON_TER	1	1
FT	PROGONADOLIBERIN I.			FT	SIGNAL	<1	5
FT	GNADOLIBERIN I.			FT	CHAIN	6	>67
FT	GNRH-ASSOCIATED PEPTIDE I (BY			FT	PEPTIDE	6	15
FT	SIMILARITY).			FT	PEPTIDE	10	15
FT	APPEARS TO BE ESSENTIAL FOR BIOLOGICAL			FT	ACT_SITE	19	>67
FT	ACTIVITY (BY SIMILARITY).			FT	MOD_RES	8	8
FT	PYRROLIDONE CARBOXYLIC ACID (BY			FT	MOD_RES	6	6
FT	SIMILARITY).			FT	AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY		
FT	AMIDATION (G-11 PROVIDE AMIDE GROUP)			FT	NON_TER	15	15
FT	SIMILARITY).			FT	SEQUENCE	67 AA;	67
FT	SIMILARITY).			FT	SEQUENCE	7573 MW;	7573 MW;
SQ	SEQUENCE	63 AA:	7370 MN;	SQ	SEQUENCE	505394DAA261A3F2 CRC64;	

Searched: 141681 seqs, 52070155 residues
RESULT 1
CON1_SHEEP

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Created: 15-DEC-1998 (Rel. 37, Last sequence update)
Last annotation update: 28-FEB-2003 (Rel. 41)

post-processing: Minimum Match 0% Maximum Match 100% Listing First 45 summaries DE Luteinizing hormone-releasing hormone I (Gonadotropin-releasing hormone I) (LH-RH I) (LH-RH-associated peptide I) DE (Fragment) DE (Fragment)

SwissProt_32:	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is dominated by one of the following:
	OS Ovis aries (Sheep)
	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
	OC Bovidae; Caprinae; Ovis.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
					DB ID	Description
1	52	96.3	61	1 GONI_SHEEP	Q98588 ovis aries	
2	52	96.3	63	1 GONI_MESAJ	O09163 mesociricetus	
3	52	96.3	67	1 GONI_MACMU	P55247 macaca mulatta	
4	52	96.3	89	1 GONI_XENELLA	P45656 xenopus laevis	
5	52	96.3	90	1 GONI_MOUSE	P23562 mus musculus	
6	52	96.3	91	1 GONI_RANCA	Q90Y63 rana catesbeiana	
7	52	96.3	91	1 GONI_PIG	P49921 sus scrofa	
8	52	96.3	92	1 GONI_HUMAN	P01148 homo sapiens	
9	52	96.3	92	1 GONI_RAT	F07490 rattus norvegicus	
10	52	96.3	92	1 GONI_TURGB	Q95335 turpia glis	
11	48	88.9	10	1 GONI_ALLMI	P37041 alligator mississippiensis	
12	48	88.9	10	1 GONI_CHICK	P37042 gallus gallus	
13	46	85.2	94	1 GONI_HABPU	P51918 haplochromis bimaculatus	
14	46	85.2	95	1 GONI_MORSA	Q73812 morone saxatilis	
15	46	85.2	95	1 GONI_PAGNA	P70074 pagrus major	
16	46	85.2	95	1 GONI_SPEAU	P51919 sparidae aurata	
17	46	85.2	99	1 GONI_DICLIA	Q9ia010 dicliduridae	
18	44	81.5	90	1 GONI_RANDY	Q9iau2 rana dybowskii	
19	43	79.6	92	1 GONI_CAYPO	P54713 carybia porcata	
20	42	77.8	80	1 GONI_CLAGA	P34349 clarias gariepinus	
21	42	77.8	91	1 GONI_ORFLA	Q9dg58 oprogonus elongatus	
22	41	75.9	10	1 GONI_CLEJPA	P81749 cleopatra pallida	
23	41	75.9	110	1 YHRJ_ACTAC	P956769 actinobacillus	
24	39	72.2	103	1 GONI_ONCKE	P55267 oncorhynchus tshawytscha	
25	39	72.2	74	1 GONI_ONCMTS	P55246 oncorhynchus mykiss	
26	39	72.2	74	1 GONI_ONCMTS	Q92097 oncorhynchus keta	
27	39	72.2	82	1 GONI_ONCMA	P30973 oncorhynchus masou	
28	39	72.2	82	1 GONI_SALUSA	P35629 salmo salar	
29	39	72.2	82	1 GONI_SALTR	P45653 salmo trutta	
30	39	72.2	89	1 GONI_PORNO	P51922 porichthys porosus	
31	39	72.2	90	1 GONI_DICLIA	Q9ia09 dicliduridae	
32	39	72.2	90	1 GONI_HABPU	P45652 haplochromis	
33	39	72.2	90	1 GONI_CORYA	P23562 coryphaenoides	

גִּנְעֹל

RESULT 1
GONI-SHEEP STANDARD; PRT; 61 AA.
 ID GONI-SHEEP
 AC Q85568;
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1995 (Rel. 37, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadotropin I precursor (Ctontopsins; Gonadotropin I; Luteinizing hormone-releasing hormone I) (Gona-
 DE hormone I) (GNRH I) (Luisiberin I); GNRH-associ-
 DE (Fragment).
 GN GNRH I OR GNRH OR LH/RH.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metzoa; Chordata; Craniata; Vertebrata;
 OC Mammalia; Butharia; Cetartiodactyla; Ruminantia;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE OF 12-61 FROM N.A.
 RC STRAIN=Western range; TISSUE=Hypothalamus;
 RA Rodriguez R.E.; Wise M.E.;
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJU d
 RN [2]
 RP SEQUENCE OF 1-10.
 RX MEDLINE:4550508;
 RA Burgess R., Butcher M., Amos M., Ling N., Monah
 RA Fellows R., Blackwell R., Vale W., Guillemin R.
 RT "Primary structure of the ovine hypothalamic lu-
 releasing factor (LRF) (LH-hypothalamus-LRF) gas
 RT spectrometry-decapeptide-Edman degradation".
 RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
 CC !-!- FUNCTION: Stimulates the secretion of gonad-
 CC the secretion of both luteinizing and follilin
 CC hormones.
 CC !-!- SUBCELLULAR LOCATION: Secreted.
 CC !-!- SIMILARITY: Belongs to the Gnrh family.
 CC This SWISS-PROT entry is copyright. It is produc-
 CC between the European Bioinformatics Institute. There are
 CC use by non profit institutions as long as it
 CC modified and this statement is not removed. Use
 CC entities requires a license agreement (See http://
 CC or send an email to license@ebi.ac.uk).
 CC EMBL: U02517; AAA3433.1; -.
 DR InterPro: IPR002012; Gnrh.
 DR InterPro: IPR004019; Gonadotropin.
 DR Pfam: PF00446; Gnrh.
 DR PRO1541; GONADOLIBRIN.
 DR PROSITE: PS00473; GNRH.
 DR Cleavage on pair of basic residues; Hormone; Ami-
 KW Placenta; Pyrrolidone carboxylic acid.
 FT NON TER 1
 FT CHAIN 1
 >61
 PROCONADOLIBERIN I
 PROCONADOLIBERIN I

CC EMBL: U00062; AAB68915.1; -
DR PIR: S46746; S46746.
DR GermOnline: 139356;
DR SGD: S0001081; MSC7.
DR GO: GO:0007133; P:meiotic recombination; IMP.
DR InterPro: IPR002086; Aldehyde_dehydr.
DR Pfam: PF00171; aldeddh; 1.
DR PROSITE: PS00070; ALDENYDE_DEHYDR_CYS; 1.
DR PROSITE: PS00087; ALDENYDE_DEHYDR_GLU; 1.
DR Hypothetical protein; Oxidoreductase.
FT ACT_SITE 354 BY SIMILARITY.
FT ACT_SITE 389 BY SIMILARITY.
SQ SEQUENCE 644 AA; 71320 MW; 5ADADAEBA16D4D CRC64;
Query Match 50.6%; Score 40; DB 1; Length 644;
Best Local Similarity 60.0%; Pred. No. 18;
Matches 9; Conservative 1; Nismatches 5; Indels 0; Gaps 0;
Qy 2 QYIKANSKFKIGITEL 16
Db : | | | | | | | |
38 QIQQDNQKLIGITL 52

Search completed: March 10, 2004, 09:13:47
Job time : 4.92607 secs

"Biochemical and genetic characterization of PepF, an oligopeptidase from Lactococcus lactis"; "Binds 1 zinc ion (By similarity).
RT J. Biol. Chem. 269:32070-32076 (1994).

[2]
RN SEQUENCE FROM N.A.
RC STRAIN=NCDO 763;
RX MEDLINE=97152670; PubMed=9209029;
RA Nardi M., Renault P., Monnet V.;
RT "Duplication of the PepF gene and shuffling of DNA fragments on the
lactic plasmid of Lactococcus lactis.";
RL J. Bacteriol. 179:1164-1171 (1997).
CC -!- FUNCTION: Hydrolyzes Peptides containing between 7 and 17 amino
 acids with a rather wide specificity.
CC -!- COFACTOR: Binds 1 zinc ion (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M3.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AB006403; AAK05825; 1; -.
DR PIR; G86840; GS6840.
DR InterPro; IPR0006025; Pept_M_Zn_BS.
DR InterPro; IPR001567; Peptidase_M3.
DR InterPro; IPR004438; Peptidase_M3B.
DR Pfam; PF01432; Peptidase_M3; 1.
DR TIGRFAMS; TIGR00181; PepF_1.
DR PROSITE; PS00142; ZINC_Protease; Zinc_CComplete_Proteome.
KW Hydrolase; Metalloprotease; Zinc (CATALYTIC) (BY SIMILARITY).
PT METAL; 387 387 ZINC (CATALYTIC) (BY SIMILARITY).
PT ACT_S-TE 388 388 ZINC (CATALYTIC) (BY SIMILARITY).
PT METAL 391 391 ZINC (CATALYTIC) (BY SIMILARITY).
PT METAL 394 394 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 601 AA; 69883 MW; 043E6B73F10FF19A CRC64; -.

Query Match 50 / 6%; Score 40; DB 1; Length 601;
Best Local Similarity 46.7%; Pred. No. 17;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
PRT :|||:|||:
Db 284 RYELRKILGTLDB 298

RESULT 15
YHJ9_YEAST STANDARD PRT; 644 AA.
ID YHJ9_YEAST
AC P38494;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical aldehyde-dehydrogenase like protein in *Filli*-YMA10
GN YR039C.
OS *Saccharomyces cerevisiae* (Baker's yeast).
OC Eukaryote; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=932; [1]
RN [1]
RP RP STRAIN=S288C / AB972;
RC MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L.W., Jier M., Johnston L., Langston Y.,
RA Lareille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Riffkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohldman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome
VIII.";
RL Science 265:2077-2082(1994).
CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC ---
RN SEQUENCE FROM N.A.
RC STRAIN=LL1403;
RX MEDLINE=11235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium *Lactococcus*
lactis ssp. *lactis* IL1403.";
RT Genome Res. 11:731-753 (2001);
CC -!- FUNCTION: Hydrolyzes peptides containing between 7 and 17 amino

AC P43013; 01-NOV-1995 (Rel. 1. 32, Created)
DT DT 01-NOV-1995 (Rel. 1. 32, Last sequence update)
DT DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable acyl carrier protein phosphodiesterase (EC 3.1.4.14) (ACP phosphodiesterase).
DE ACPD OR HU366.
GN Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
RN [1] NCBI_TaxID=727;

RP SEQUENCE FROM N.A.
RN STRAIN=ATCC 27343 / Kid;
RN RC STRAIN=ATCC 27343 / Kid;
RN RX MEDLINE=88142549; PubMed=1481422;
RN RA Okunbo S., Muto A., Kawachi Y., Yamao F., Osawa S.;
RN RT "the ribosomal protein gene cluster of Mycoplasma capricolum.";
RN RL Mol. Genet. 210:314-322 (1987).
RN CC -!
RN CC -! FUNCTION: This small ubiquitous enzyme is essential for maintenance and cell growth.
RN CC -! CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
RN CC -! SUBUNIT: Monomer (By similarity).
RN CC -! SUBCELLULAR LOCATION: Cytoplasmic.
RN CC -! SIMILARITY: Belongs to the adenylate kinase family.

CC -!
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

RP SEQUENCE FROM N.A.
RN STRAIN=Rd / KW20- / ATCC 51907;
RN RX MEDLINE=95350630; PubMed=542880;
RN RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkinness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., FitzHugh W., Fields C.A., Gocayne J.D., Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley R., Utterback T.R., Phillips C.A., Spriggs T., Heidelberg J.M., Fine L.D., Fritchman J.L., Nguyen D.T., Saudek D.M., Brandon R.C., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;
RN CC "Whole genome random sequencing and assembly of Haemophilus influenzae Science 269:496-512(1995)."
RN CC -! FUNCTION: Converts holo-ACP to apo-ACP by hydrolytic cleavage of the phosphopantetheine prosthetic group from ACP (By similarity).
RN CC -! CATALYTIC ACTIVITY: Holo-lacyl-carrier protein + H(2)O = 4'-phosphopantetheine + apo-lacyl-carrier protein].
RN CC -! SIMILARITY: Belongs to the acpD family.

CC -!
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

DR EMBL; U20944; ARNC43728.1; -;
DR PIR; C64026; AAC23013.1; -;
DR TIGR; HU1366; -;
DR InterPro; IPR001680; NADHdh_2;
DR Pfam; PF0225; Flavodoxin_2; 1.
SQ SEQUENCE 194 AA; Complete protein. Hydroxylase; 21208 MW; A8AECD00829522C CRC64;
SQ Score 40; DB 1; Length 194;
Best Local Similarity 53.3%; Pred. No. 5.4;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 QVIKANSKFIGITEL 16
Db 147 QYMKSLIIGFIGITDV 161

RESULT 13
PEPF1_LACLC STANDARD; PRT: 601 AA.
ID PEP1_LACLC
AC P34124 / P94880;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oligoendorpeptidase F, Plasmid (EC 3.4.24.-).
GN PEPF1 OR PEPF.
OS Lactococcus lactis (subsp. cremoris) (Streptococcaceae; Lactococcus).
OC Plasmid pL2763;
RN [1] NCBI_TaxID=1359;
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RAD_MYCCA
ID RAD_MYCCA
AC P10251;
DT 01-MAR-1989 (Rel. 1. 10, Created)

RESULT 12
RAD_MYCCA STANDARD; PRT: 213 AA.
ID RAD_MYCCA
AC P10251;
DT 01-MAR-1989 (Rel. 1. 10, Created)

RX MEDLINE=9509644; PubMed=7798200;
RA Monnet V., Nardi M., Chopin A., Chopin M.-C., Gripon J.-C.;

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AP174546; AAD50307.1; ALT_INIT
HANAP; MP_00219; -; 1.
InterPro; IPR016580; Amidohydro_1.
DR InterPro; IPR00215; Pept_NMB.nph.
DR Pfam; PF001979; Amidohydro_1; 1.
PROSITE; PS00489; DIHYDROPROTASE_1; PARTIAL.
DR PROSITE; PS00483; DIHYDROPROTASE_2; 1.
KW Pyrimidine biosynthesis; Hydrolase; Zinc.
FT NON_TER 1 1
FT ZINC 2 1 BY SIMILARITY.
FT METAL 34 34 ZINC 2 1 BY SIMILARITY.
FT METAL 107 107 ZINC 1 1 BY SIMILARITY.
SEQUENCE 204 AA; 22899 MW; 9DB0B6C9B834B310 CRC4;

Query Match Score 41; DB 1; Length 204;
Best Local Similarity 51.9%; Pred. No. 3/7;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
Qy 2 QYIKANSKFIGIT 14
Db 42 QYVQAGNRFLGAT 54

RESULT 9

LEI12 THETN STANDARD; PRT; 384 AA.
AC Q8RCF9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 2-isopropylmalate Synthase 2 (EC 2.3.3.13) (Alpha-isopropylmalate synthase 2) (Alpha-IPM synthetase 2).
GN LEUA2 OR TTE0472.
RN [1] CC Thermanaerobacter tengcongensis; Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
RN [1] CC STRAIN=MB4 / JCM 11007;
RN MEDLINE=21:992816; PubMed=11997336;
RN Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J., Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L., Tan H., Chen R., Wang J., Yu J., Yang H.;
RN A complete sequence of *T. tengcongensis* genome.";
RN Genome Res. 12:689-700 (2002).

-|- FUNCTION: Catalyzes the condensation of the acetyl group of acetyl-CoA with 3-methyl-1,2-oxobutanoate (2-isopropylmalate) to form 3-carboxy-3-hydroxy-4-methylipenanoate (2-isopropylmalate).
CC -|- CATALYTIC ACTIVITY: Acetyl-CoA + 3-methyl-1,2-oxobutanoate + H(2)O = 2-hydroxy-2-isopropylsuccinate + CoA.
CC -|- PATHWAY: Leucine biosynthesis; first step.
CC -|- SUBUNIT: Homotrimer (By similarity).
CC -|- SIMILARITY: Belongs to the alpha-IPM synthetase / homocitrate synthase family. LeuA 1 subfamily.
CC DR EMBL; AE013018; AAC23753.1; -.
CC DR HANAP; MF_01025; atypical; 1.
CC DR InterPro; IPR002034; AIPM/Hcit_synth.
CC DR InterPro; IPR0010891; HMGL-like.
CC DR PFO0682; HMGL-like; 1.

PROSITE; PS00815; AIPM_HOMOCIT_SYNTH_1; 1.
PROSITE; PS00816; AIPM_HOMOCIT_SYNTH_2; 1.
Leucine biosynthesis; Trasferase; Complete proteome.
SEQUENCE 384 AA; 4240 MW; 095310F2C0E4ADD CRC64;

Query Match Score 41; DB 1; Length 384;
Best Local Similarity 51.9%; Pred. No. 7/1;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Qy 4 IKANSKFIGIT 14
Db 214 VKAGAKKFVGT 224

RESULT 10

V120_HSVSA STANDARD; PRT; 899 AA.
ID V120_HSVSA
AC Q01055;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Capsid assembly protein 63.
DN 63 OR EEBP1.
OS Herpesvirus saimiri (strain 11).
OC dsDNA viruses; no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10383;
RN [1] CC SEQUENCE FROM N.A.
RN MEDLINE=9233388; PubMed=11321287;
RN RX ALbrecht J.-C., Nicholas J., Billier D., Cameron K.R., Biesinger B., Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B., Honess R.W.;
RN "Primary structure of the herpesvirus saimiri genome.";
RN RL J. Virol. 66:5047-5058 (1992).
RN [2] CC SEQUENCE FROM N.A.
RN MEDLINE=9230228; PubMed=11314457;
RN RX Nicholas J., Cameron K.R., Coleman H., Newman C., Honess R.W.;
RN RT "Analysis of nucleotide sequence of the rightmost 43 kbp of herpesvirus saimiri (HVS) L-DNA: General conservation of genetic organization between HVS and Epstein-Barr virus.";
RN RL J. Virol. 168:296-310 (1992).
RN [-] CC SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL37, EH-1, 23, BEV BOLF1, VZV 21, HVS-1 63, AND HCMV UL47.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

SEQUENCE FROM N.A.

MEDLINE=9233388; PubMed=11321287;

RX ALbrecht J.-C., Nicholas J., Billier D., Cameron K.R., Biesinger B.,

Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,

Honess R.W.;

"Primary structure of the herpesvirus saimiri genome.";

RN RL J. Virol. 66:5047-5058 (1992).
RN [2] CC SEQUENCE FROM N.A.

RN MEDLINE=9230228; PubMed=11314457;

RX Nicholas J., Cameron K.R., Coleman H., Newman C., Honess R.W.;

RT "Analysis of nucleotide sequence of the rightmost 43 kbp of

herpesvirus saimiri (HVS) L-DNA: General conservation of genetic

organization between HVS and Epstein-Barr virus.";

RN RL J. Virol. 168:296-310 (1992).
RN [-] CC SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL37, EH-1, 23, BEV BOLF1, VZV 21, HVS-1 63, AND HCMV UL47.CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

SEQUENCE 899 AA; 103350 MW; F1429B3770A2885B CRC64;

DR X64346; CR45686.1; -.

DR M86405; ARA46139.1; -.

DR InterPro; IPR008643; Herpes_ORF63.

DR PFO5755; Herpes_ORF63; 1.

KW Capsid assembly.

SEQUENCE 899 AA; 103350 MW; F1429B3770A2885B CRC64;

Query Match Score 41; DB 1; Length 899;

Best Local Similarity 51.9%; Pred. No. 17;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 QYTRANSKFIGITE 15
Db 124 QYITSNATPTGISE 137

RESULT 11
ACPD_HAEIN STANDARD; PRT; 194 AA.
ID ACPD_HAEIN
ID ACPD_HAEIN

RA Weber J.L., Sim B.K.L., Lyon J.A., Wolff R.;
 RT "Merozoite surface protein sequence from the Camp strain of the human
 RL Malaria parasite Plasmodium falciparum.";
 RC Nucleic Acids Res. 16:1206-1206(1988).
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC -!- (Potential).
 CC Merozoite surface antigen contain the sequence of 83 kDa, 42
 CC kDa and 19 kDa antigens which are the major surface antigens of
 CC merozoites. The maturation take place during schizont.
 CC -!- (Potential). It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -!- (Potential).
 DR EMBL; X03831; CAA27446.1; -.
 DR PIR; A2A316; SA2ZQGM.
 DR InterPro; IPR006209; EGF_1-like.
 DR Pfam; PF00008; EGF_1.
 DR Malaria; Merozoite; Polypeptide; Repeat; Glycoprotein;
 KW Transmembrane; GPI-anchor.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 1726 MEROZOITE SURFACE PROTEIN 1.
 PT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 827 827 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 924 924 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 944 944 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 990 990 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 1016 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 1114 1114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 1221 1221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 PT CARBOHYD 1613 1613 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1726 AA; 195197 MW; S559CBEFA229A026 CRC44;
 FT CARBOHYD 944 944 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 990 990 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1221 1221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1613 1613 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1726 AA; 195197 MW; DD8AD5FA52BCF3 CRG44;
 Query Match 53.8%; Score 42.5%; DB 1; Length 1726;
 Best Local Similarity 60.0%; Pred. No. 17;
 Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

Qy 2 QYIKANSKPI-GITE 15
 Db 1026 QFVKNSNSKVIGLT 1040

RESULT 8 PYRC_SPERMA STANDARD; PRT; 204 AA.
 ID PYRC_SPERMA
 AC Q933S1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dihydrocoenzyme (EC 3.5.2.3) (DHOase) (Fragment).
 GN PYRC.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Serratia.
 OC NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=SM6;
 RA Berkman M.; Benedik M.J.;
 RC "Dihydrocoenzyme transcription of Serratia marcescens nuclease.";
 RC Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: (S)-dihydrocoerate + H(2)O = N-carbamoyl-L-
 CC -!- ASPARTATE; Binds 2 zinc ions per subunit (By similarity).
 CC -!- COACTOR: Pyrimidine biosynthesis; third step.
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SIMILARITY: Belongs to the DHCase family. Subfamily 1.
 CC -!- (Potential). It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor

CC	-1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (Potential)	RA Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U., Stunnenberg H., Bujard H.; "Proteomorphism of the Precursor for the major surface antigens of Plasmodium falciparum merozoites: studies at the genetic level.", J. Mol. Biol. 1985; 188(1):383-392.
CC	-1- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42 kDa and 19 kDa antigens which are the major surface antigens of merozoites. The maturation take place during schizont.	CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (Potential).
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	CC -1- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42 kDa and 19 kDa antigens which are the major surface antigens of merozoites. The maturation take place during schizont.
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	CC -1- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42 kDa and 19 kDa antigens which are the major surface antigens of merozoites. The maturation take place during schizont.
DR	EMBL; M19143; AAA29653.1; -	DR EMBL; X05624; CAa29112.1; -
DR	InterPro; IPR006309; EGF-like	DR InterPro; IPR006209; EGF-like
DR	Pfam; PF00038; EGF; 1;	DR Pfam; PF00038; EGF; 1;
KW	Malaria; Merozoite; Polypeptide; Glycoprotein;	KW Malaria; Merozoite; Polypeptide; Glycoprotein;
KW	Transmembrane; GPI-anchor.	KW Transmembrane; GPI-anchor.
FT	SIGNAL 1	FT SIGNAL 1
FT	CHAIN 19	FT CHAIN 19
FT	CARBOHYD 20 1701 MEROZOITE SURFACE PROTEIN 1.	FT CARBOHYD 20 1701 MEROZOITE SURFACE PROTEIN 1.
FT	CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 802 802 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 802 802 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 899 899 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 899 899 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 919 919 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 919 919 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 965 965 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 965 965 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 991 991 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 991 991 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1089 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 1089 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1196 1196 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 1196 1196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1588 1588 N-LINKED (GLCNAC. . .) (POTENTIAL).	FT CARBOHYD 1588 1588 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQUENCE	1701 AA; 193719 MW; 390B75E73D3B552 CRC64;	SEQUENCE 1701 AA; 193768 MW; 3FC2EC5PAF963A98 CRC64;
Qy	Query Match Best Local Similarity 53.8%; Score 42.5%; DB 1; Length 1701; Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;	Qy 2 QYIKANSKPI-GITE 15
Db	2 QYIKANSKPI-GITE 15 1001 QYIKANSKPI-GITE 1015	Db 1001 QYIKANSKPI-GITE 15
RESULT 5		RESULT 6
MSPL_PLAFM	MSPL_PLAFM STANDARD; PRT; 1701 AA.	MSPL_PLAFM STANDARD; PRT; 1726 AA.
ID	MSPL_PLAFM	ID MSPL_PLAFM
AC	F085639; DT 01-AUG-1988 (Rel. 08, Created)	AC P04934; DT 13-AUG-1987 (Rel. 05, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)	DT 01-MAR-1989 (Rel. 10, Last sequence update)
DE	Merozoite surface protein 1 precursor (Merozoite surface antigens) (PMMSA) (P190).	DT 15-MAR-2004 (Rel. 43, Last annotation update)
DB	OS Plasmodium falciparum (isolate mad20 / Papua New Guinea); Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium; NCBI_TaxID=70153;	DE Merozoite surface protein 1 precursor (Merozoite surface antigens) (PMMSA) (P195).
RN	[1]	DN MSP-1.
RP	SEQUENCE FROM N.A. MEDLINE-88011243; PubMed=3079521; Tanabe K., Mackay M., Goman M., Scaife J.G.; "Allelic dimorphism in a surface antigen gene of the malaria parasite Plasmodium falciparum"; J. Mol. Biol. 1987; 193:273-287(1987).	OS Plasmodium falciparum (isolate Camp / Malaysia); Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN	[2]	OC NCBITaxID=5835;
RR	REVISIONS TO 1403; 1569 AND 1625.	RN SEQUENCE OF 1-1103 FROM N.A.
RN	Tanabe K.; Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.	RX MEDLINE=86205236; PubMed=517809;
RN	[3]	RN Weber J.L., Leininger W.M., Lyon J.A.; "Variation in the gene encoding a major merozoite surface antigen of the human malaria parasite Plasmodium falciparum." Nucleic Acids Res. 14:3311-3323(1986).
RP	SEQUENCE OF 1-115 FROM N.A. MEDLINE=86136024; PubMed=3004972;	RX SEQUENCE OF 1104-1726 FROM N.A. MEDLINE=88143999; PubMed=3278296;

Query Match Score 79; DB 1; Length 1314;
 Best Local Similarity 100.0%; Pred. No. 2.7e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OX [1] 1 MOYIKANSKFIGITEL 16
 Db 828 MOYIKANSKFIGITEL 843

RESULT 2

VG84_BPM15 STANDARD; PRT; 66 AA.
 ID -VG84_BPM15
 AC Q05301; Q255921;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Gene 84 protein (GB84).
 GN 84.
 OS Mycobacteriophage L5.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
 OC L5-like viruses.
 OX NCBI_TAXID=31757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=9311282; PubMed=8459766;

RA Hattori G.F., Sarkis G.J.;
 RT "DNA sequence, structure and gene expression of mycobacteriophage L5:
 RT a phage system for mycobacterial genetics.";
 RL Mol. Microbiol. 7:355-405(1993).
 CC

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions. Usage as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC

DR EMBL; 218946; CA879460.1; -.

DR PIR; S31029; S31029; 7424 MW;

SQ SEQUENCE 66 AA; 7424 MW; 9C7104C7A4FA74AS CRC64;

Query Match Score 44; DB 1; Length 66;
 Best Local Similarity 55.7%; Pred. No. 0.33%;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 YIKANSKFIGITEL 16
 Db 50 YIKANGKFIGITEL 63

RESULT 3

MSP1_PLAF3 PLAF3 STANDARD; PRT; 1682 AA.
 ID MSP1_PLAF3
 AC P19579; Q255921;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Plasmodium falciparum 1 precursor (Merozoite surface antigens)
 DE (PFMMSA) (P190).
 GN MSP-1.

OS Plasmodium falciparum (isolate ro-33 / Ghana); Haemosporida; Plasmodium;
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OC
 OX NCBI_TAXID=5834;

RN [1]

RP SEQUENCE OF 1-1061 FROM N.A.

RX MEDLINE=8816657; Published=3-27-688;

RA Cetera U., Rotman D., Matile H., Reber-Liske R.;

"A naturally occurring gene encoding the major surface antigen

precursor P190 of Plasmodium falciparum lacks tripeptide repeats."

RL Embio J. 6:4137-4142(1987).

[2]

SEQUENCE OF 1032-1682 FROM N.A.
 RX MEDLINE=9535493; PubMed=7628566;

RA Toile P., Buillard H., Cooper J.A.;
 RT "Plasmodium falciparum: variations within the C-terminal region of
 merozoite surface antigen-1.";

RL Exp. Parasitol. 81:447-54(1998).

CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor

CC (Potential).

CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42

CC kDa and 19 kDa antigens which are the major surface antigens of

CC merozoites. The maturation take place during schizont.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions. Usage as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

or send an email to license@isb-sib.ch).

CC DR EMBL; M35727; AAA29715.1; -.

DR EMBL; Y0087; CAA68380.1; -.

DR EMBL; 235326; CAA8455.1; -.

DR InterPro; IPR006209; EGF-like.

DR Pfam; PF0008; EGF_1.

KW Malaria; Merozoite; Polypeptide; Repeat; Signal; Glycoprotein;

KW Transmembrane; GPI-anchor.

POTENTIAL.
 FT SIGNAL 1 19

FT CHAIN 1 20

FT TRANSMEM 1666 1682

FT CARBOHYD 233 233

FT CARBOHYD 462 462

FT CARBOHYD 528 528

FT CARBOHYD 599 599

FT CARBOHYD 785 785

FT CARBOHYD 881 881

FT CARBOHYD 901 901

FT CARBOHYD 947 947

FT CARBOHYD 1071 1071

FT CARBOHYD 1178 1178

FT CARBOHYD 1569 1559

FT N-LINKED (GLCNAC, -)

Eur. J. Biochem. 188:39-45 (1990).
 [6] PARTIAL SEQUENCE. MEDLINE=92037649; PubMed=1935979;
 RX Kriegstein K.G.; Henschien A.H.; Weller U.; Habermann E.;
 RA "limited proteolysis of tetanus toxin. Relation to activity and
 identification of cleavage sites."
 RT Eur. J. Biochem. 202:41-51 (1991).
 RL RN [7]
 RP IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93010948; PubMed=1331607;
 RA Schiavo G., Poulaïn B., Rossetto O., Benfenati F., Tauc L.,
 RA Montecucco C.; Dasquata B.R., Montecucco C.;
 RT "Tetanus toxin is a zinc protein and its inhibition of
 neurotransmitter release and protease activity depend on zinc.";
 RL EMBO J. 11:3577-3583 (1992).
 RN RN [8]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331607;
 RA Schiavo G., Benfenati F., Poulaïn B., Rossetto O., de Laureto P.P.,
 RA Dasquata B.R., Montecucco C.;
 RT RT by proteolytic cleavage of synaptobrevin."
 RL Nature 359:832-835 (1992).
 RN RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
 RX MEDLINE=97475217; PubMed=9334741;
 RA Umland T.C., Wigert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
 RA Sax M.;
 RT "Structure of the receptor binding fragment HC of tetanus
 neurotoxin."
 RL Nat. Struct. Biol. 4:788-792 (1997).
 CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL TRANSMITTERS, IS INTERNALIZED
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN- | -PHE-77
 BOND OF SYNAPTOBREVIN-2.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln- | -Phe-77 bond in
 CC synaptobrevin-2.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (BY SIMILARITY).
 CC -!- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
 CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
 CC AND ARE NON TOXIC AFTER SEPARATION.
 CC -!- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
 CC GRANULOZYME RECEPTORS.
 CC -!- SIMILARITY: Belongs to peptidase family M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X04436; CAA28033.1;
 DR EMBL; X06214; CAA29564.1;
 DR EMBL; AF528097; AAO31454.1;
 DR PIR; A25689; BTGLTN;
 DR PDB; 1AF9; 29-APR-98.
 DR PDB; 1ABD; 14-OCT-98.
 DR PDB; 1D0H; 27-MAR-00.
 DR PDB; 1DFQ; 24-MAR-00.
 DR PDB; 1DIW; 24-MAR-00.
 DR PDB; 1DII; 14-MAR-00.
 DR PDB; 1FV3; 05-SEP-01.
 DR MEROPS; M27_001;
 DR InterPro; IPR000985; ConX-like_lec_g1.
 DR InterPro; IPR00160; Kunz_x_legume.
 DR InterPro; IPR00025; Pept_M_Zn_BS.

Scoring table:	BLOSSUM62	34	37	46.8	757	1	245D YEAST
Searched:	Gapop 10.0 , Gapext 0.5	35	37	46.8	774	1	RBP3-TNCBE
Total number of hits satisfying chosen parameters:	141681 seqs, 52070155 residues	36	37	46.8	774	1	RFP3-TINBJU
Post-processing:	Minimum Match 0%	37	36	45.6	169	1	Y358_BUGAI
Database :	SwissProt_42.*	38	36	45.6	247	1	Q91577_xenopus_lae
Scored by:	US-09-848-834A-2	39	36	45.6	258	1	C561_XENLA
Sequence:	1 MQYIKANSKFIGITTEL 16	40	36	45.6	296	1	Q977P3_chlamydia_p
		41	36	45.6	329	1	P33659_clostridium
		42	36	45.6	333	1	O57383_NADAL_CYAPA
		43	36	45.6	461	1	DPOB_XENLA
		44	36	45.6	495	1	NIFN_RHOCA
		45	36	45.6	630	1	P11410_pichia_jadi
							P40009_saccharomyces_cerevisiae_YND1_YEAST
ALIGNMENTS							
RESULT 1							
ID	TETX_CLOTE	AC	P04958;	STANDARD;	PRT;	1314 AA.	
DT	13-AUG-1987 (Ref. 05, Created)	DT	13-AUG-1987 (Ref. 05, Last sequence update)				
DB	Tetanus toxin Precursor (EC 3.4.24.68) (Tetatoxylisin) [Contains: Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy chain (Tetanus toxin chain H)].	DT	10-OCT-03 (Ref. 42, Last annotation update)				
DE	Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy chain (Tetanus toxin chain H).	DE					
GN	TETBX OR CTP60.	GN					
OS	Clostridium tetani.	OS					
OG	Plasmid pPE8 and Plasmid 75 Kbp; Bacteria; Firmicutes; Clostridia; Clostridiaceae;	OG					
OC	Clostridium.	OC					
NCBI_TaxID=1513;	NCBI_TaxID=1513;	OX					
SEQUENCE FROM N.A.							
RC	PLASMID=75 Kbp;	RC	PLASMID=75 Kbp; MEDLINE=87053814; PubMed=3536478;				
RA	Eisel U., Jarausch W., Goeretzai K., Henschen A., Engel G. J., Weller U., Hudek M., Habermann E., Niemann H.;	RA					
RT	"Tetanus toxin: primary structure, expression in <i>E. coli</i> , and homology with botulinum toxins."	RT					
RL	EMBO J. 5:2495-2502 (1986).	RL					
SEQUENCE FROM N.A.							
RC	PLASMID=75 Kbp; PLASMID=75 Kbp; MEDLINE=87040747; PubMed=3775474;	RC					
RA	Fairweather N.F., Lyness V.A.;	RA					
RT	"The complete nucleotide sequence of tetanus toxin."	RT					
RL	Nucleic Acids Res. 14:7809-7812 (1986).	RL					
SEQUENCE FROM N.A.							
RC	SEQUENCE FROM N.A. STRAIN=Massachusetts / 388; PLASMID=pPE88; MEDLINE=2245723; PubMed=12552129;	RC					
RA	Brueggemann H., Baumer S., Pricke W.F., Wiezer A., Liesegang H., Decker T., Herberg C., Martinez-Arias R., Merkl R., Henne A., Gottschalk G.;	RA					
RT	"The genome sequence of Clostridium tetani, the causative agent of tetanus disease."	RT					
RL	Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321 (2003). [4]	RL					
SEQUENCE OF 742-1314 FROM N.A.							
RC	PLASMID=75 Kbp;	RC					
RA	Brueggemann H., Baumer S., Pricke W.F., Wiezer A., Liesegang H., Decker T., Herberg C., Martinez-Arias R., Merkl R., Henne A., Gottschalk G.;	RA					
RT	"Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in <i>Escherichia coli</i> ."	RT					
RL	RT						
PARTIAL SEQUENCE, AND DISULFIDE BONDS.							
RC	PARTIAL SEQUENCE, AND DISULFIDE BONDS. MEDLINE=90201034; PubMed=2105021;	RC					
RA	Kriegstein K., Hensch A., Wellier U., Habermann E.;	RA					
RT	"Arrangement of disulfide bridges and positions of sulphydryl groups in tetanus toxin."	RT					
RL							

Wed Mar 10 10:34:30 2004

usb-09-848-834a-2.open.rpr

Page 5

Qy 3 YIKANSKFIGITTEL 16
Db 57 YIKYNEKEFGITTEL 70

Search completed: March 10, 2004, 09:16:40
Job time : 7.03937 secs

major merozoite surface antigen Precursor - malaria parasite (Plasmodium falciparum) (strain O157:H7, substrain EDL93) submitted to the EMBL Data Library, November 1993
 N; Alternate names: 195K glycoprotein
 C; Species: Plasmodium falciparum
 C; Date: 30-Sep-1987 #sequence_revision 31-Mar-1991 #text_change 09-Jun-2000
 A; Accession: A23386; S0361
 R; Weber, J.L.; Lininger, W.M.; Lyon, J.A.
Nucleic Acids Res. 14, 3311-3323, 1986
 A; Title: Variation in the gene encoding a major merozoite surface antigen of the human merozoite. Reference number: A23386; MUID:5620536; PMID:3517809
 A; Accession: A23386
 A; Residues: 1-1104 <WEBI>
 A; Cross-references: EMBL:X03831
Nucleic Acids Res. 16, 1206, 1988
 A; Title: Merozoite surface protein sequence from the Camp strain of the human malaria parasite. Reference number: S0361; MUID:8814399; PMID:3278296
 A; Accession: S06361
 A; Molecule type: DNA
 A; Residues: 1-104-1726 <WEBB2>
 A; Cross-references: EMBL:X03831
 C; Comment: The merozoite stages of different strains have strain-specific surface antigens. The merozoite, sporozoite, and gameteocyte. The merozoite surface antigen has three stages: sporozoite, merozoite, and gameteocyte.
 C; Superfamily: major merozoite surface antigen
 C; Keywords: Glycoprotein, merozoite; surface antigen; tandem repeat
 F; 1-19/Domain: signal sequence predicted <SIG>
 F; 20-176/Product: major merozoite surface antigen #status predicted <MT>
 F; 67-87, 91-96, 100-105, 109-120/Region: 3-residue repeats (S-G-T)
 F; 757-765/Region: 3-residue repeats (T-E-E)
 F; 1133-272, 501-567, 638-827, 839-924, 944-990, 1016-1114, 1221-1613, 1658/Binding site: carbohydrate
 Query Match Score 53.8%; DB 1; Length 1726;
 Best Local Similarity 60.0%; Pred. No. 42;
 Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;
 QY 2 QYIKANSKFI-GITE 15
 DB 1026 QFVKSNSKVIGLTE 1040

RESULT 12
 A45948
 Major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (strain O157:H7, substrain EDL93) submitted to the EMBL Data Library, November 1993 #sequence_revision 30-Sep-1993 #text_change 09-Jun-2000
 C; Species: Plasmodium falciparum
 C; Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jun-2000
 C; Accession: A45948
 R; Chang, S.P.; Kramer, K.J.; Yamaga, K.M.; Kato, A.; Case, S.E.; Siddiqui, W.A.
Exp. Parasitol. 67, 1-11, 1988
 A; Title: Plasmodium falciparum: gene structure and hydrophathy profile of the major merozoite surface antigen. Reference number: A45948; MUID:89005525; PMID:3049134
 A; Accession: A45948
 A; Status: Preliminary
 A; Molecule type: DNA
 A; Residues: 1-1126 <CHA>
 A; Cross-references: GB:M7213
 C; Superfamily: major merozoite surface antigen
 C; Keywords: surface antigen
 Query Match Score 53.8%; DB 2; Length 1726;
 Best Local Similarity 60.0%; Pred. No. 42;
 Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;
 QY 2 QYIKANSKFI-GITE 15
 DB 1026 QFVKSNSKVIGLTE 1040

RESULT 13
 S38635
 blastospore polyprotein - fruit fly (*Drosophila melanogaster*)
 C; Species: *Drosophila melanogaster*
 C; Date: 06-Jun-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000
 C; Accession: S38635

R; Frommer, G.; Schuh, R.; Jddkle, H.
 Submitted to the EMBL Data Library, November 1993
 A; Description: Locomotion of a novel microtia-like element in the blastoderm
 A; Reference number: S38635
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-1333 <FRO>
 A; Cross-references: EMBL:227119; NID:9415797; PID:9415798
 C; Genetics:
 A; Gene: FlyBase:microtia
 A; Cross-references: FlyBase:microtia
 C; Keywords: polyprotein
 Query Match Score 53.2%; DB 2; Length 1333;
 Best Local Similarity 53.3%; Pred. No. 40;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 2 QYIKANSKFI-GITE 16
 DB 127 KVVKQARSKMIGSAEL 141

RESULT 14
 D85794
 hypothetical protein Z2873 [imported] - Escherichia coli (strain O157:H7, substrain EDL93) submitted to the EMBL Data Library, November 1993
 C; Species: Escherichia coli
 C; Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C; Accession: D85794
 R; Perna, N.T.; Plunkett III, G.; Burl, V.; Mau, B.; Glaeser, J.D.; Rose, D.J.; Mayhew, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamitis, K.; Apodaca, Nature 409, 529-533, 2001
 A; Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A; Reference number: A85450; MUID:2107435; PMID:11206551
 A; Accession: D85794
 A; Status: Preliminary
 A; Molecule type: DNA
 A; Residues: 1-79 <STC>
 A; Cross-references: GB:AB005174; NID:912515873; PID:NAG5616.1; GSPDB:GN00145; UMG:Z;
 A; Experimental source: strain O157:H7, substrain EDL93
 C; Genetics:
 A; Gene: 22873

Query Match Score 51.9%; DB 2; Length 79;
 Best Local Similarity 46.7%; Pred. No. 36;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MOYIKANSKFI-GITE 15
 DB 51 VOYLHTRSGFLGTD 65

RESULT 15
 G48677
 Ig heavy chain V-D-J region (419.1) - mouse (fragment)
 C; Species: Mus musculus (house mouse)
 C; Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
 C; Accession: G48677
 R; Bassagoun, J.; Brait, M.; Jamila, I.; Urbain, J.; Gotlib, P.; Brown, A.; Hasemann, Proc. Natl. Acad. Sci. U.S.A. 90, 9508-9512, 1993
 A; Title: Molecular characterization of monoclonal CRI-A-positive anti-arsonate antibody. A; Reference number: A48677; MUID:9402404; PMID:8415731
 A; Accession: G48677
 A; Status: Preliminary; not compared with conceptual translation
 A; Molecule type: mRNA
 A; Residues: 1-123 <TA>
 C; Superfamily: immunoglobulin V region; immunoglobulin
 C; Keywords: heterotetramer; immunoglobulin F; 15-98: Domain: immunoglobulin homology <IMM>

Query Match Score 51.9%; DB 2; Length 123;
 Best Local Similarity 64.3%; Pred. No. 56;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

A;Residues: 1-1086 <OLA>
 A;Cross-references: EMBL:X61930
 C;Species: Melanochromis auratus
 C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 05-Nov-1999
 C;Accession: I50966; 15083
 C;Organism: H. r. Huiguen, C.; Richy, H.; Klein, J.
 C;Title: Major-histocompatibility-complex variation in two species of cichlid fishes from Mol. Biol. Evol. 10, 1060-1072, 1993
 A;Reference number: A48209; MUID:9408588; PMID:8112649
 A;Accession: I50966
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-82 <ON2>
 A;Cross-references: GB:L117482; NID:9308975; PIDN:AAA49349.1; PID:g308976
 A;Gene: 150963
 A;Cross-references: GB:L117456; NID:9309011; PID:g309012
 C;Genetics:
 C;Superfamily: class II histocompatibility antigen; immunoglobulin homology
 C;Species: Plasmodium falciparum

Query Match Score 53.8%; Score 42.5%; DB 2; Length 82;
 Best Local Similarity 47.4%; Pred. No. 2;
 Matches 9; Conservative 5; Mismatches 2; Indels 3; Gaps 1;

Qy 1 :|:|:|:|:|:|:|:|
 Db 31 MEYIRPDSVGKEYGFTQL 49

RESULT 7
 S06286 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (fragment)
 N;Alternate names: 190K protein; polymorphic schizont antigen
 C;Species: Plasmodium falciparum
 C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jun-2000
 C;Accession: S06286
 R;Certa, U.; Rotmann, D.; Matile, H.; Reber-Liske, R.
 EMBO J. 6, 4137-4142, 1987
 A;Title: A naturally occurring gene encoding the major surface antigen precursor p190 of
 A;Reference number: S06286; MUID:88166657; PMID:3327688
 A;Accession: S06286
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 C;Superfamily: major merozoite surface antigen
 C;Keywords: surface antigen

Query Match Score 53.8%; Score 42.5%; DB 2; Length 1060;
 Best Local Similarity 60.0%; Pred. No. 26;
 Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

Qy 2 QYIKANSKFI-GITE 15
 Db 983 QFVKNSKVITGLTE 997

RESULT 8
 S16752 major merozoite surface antigen MSA-1 - malaria parasite (Plasmodium falciparum) (fragment)
 N;Alternate names: Polymorphic schizont antigen p190
 C;Species: Plasmodium falciparum
 C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jun-2000
 C;Accession: A44865; S16752
 R;Olasson, P.; Matile, H.; Certo, U.
 Exp. Parasitol. 74, 381-389, 1992
 A;Title: Plasmodium falciparum: the repetitive MSA-1 surface protein of the RO-71 isolate
 A;Reference number: A44865; MUID:92275047; PMID:1523091
 A;Accession: A44865
 A;Status: preliminary
 A;Molecule type: DNA
 SAZQGM

Query Match Score 53.8%; Score 42.5%; DB 2; Length 1086;
 Best Local Similarity 60.0%; Pred. No. 26;
 Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

Qy 2 QYIKANSKFI-GITE 15
 Db 1001 QFVKNSKVITGLTE 1015

RESULT 11
 SAZQGM

R;de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
 Eur. J. Biochem. 229, 61-69, 1995

A;Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
 A;Reference number: S63348; PMID:9526688; MJD:7744050

A;Accession: S69348
 A;Molecule type: protein

A;Residues: 2-31 <DEF>
 C;Comment: The source of this protein was an extrachromosomal plasmid.

C;Comment: The precursor is cleaved by endogenous proteinase activity to form light (fragment A) and heavy (fragment B) dual chains. The amino end of the heavy chain (fragment B) binds to gangliosides in a lipid bilayer. Fragment C binds to peripheral neuronal synapses, is internalized and cleaved by proteolytic enzymes. It inhibits neurotransmitter release by blocking synaptic transmission.

C;Function:
 A;Description: blocks neurotranscytosis via hydrolysis of a Gln-Phe peptide bond in synaptic terminals.

C;Superfamily: tetanus toxin hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc

C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc

F;2-457/1;Product: tenteroxylisin light chain (Fragment A) #status Predicted <TRU>

F;461-1315/1;Product: tentoxylisin heavy chain (Fragment B) #status Experimental <TRH>

F;805-1315/1;Domain: channel forming (fragment B) #status Predicted <TRX>

F;233-237/1;Domain: ganglioside binding (fragment C) #status Predicted <TRC>

F;234-237/1;Binding site: zinc (HS) #status Predicted

F;234/Active site: Glu #status Predicted

Query Match Score 79; DB 1; Length 1315;

Best Local Similarity 100.0%; Pred. No. 8.5e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQYIKANSKFIGITEL 16
 Db 829 MQYIKANSKFIGITEL 844

RESULT 2
 S29982
 class II histocompatibility antigen - Atlantic salmon

C;Species: Salmo salar (Atlantic salmon)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999

C;Accession: S29982
 R;Hardrik, I.
 Submitted to the EMBL Data Library, October 1992

A;Accession number: S29980
 A;Status: Preliminary

A;Molecule type: mRNA
 A;Residue: 1-244 <HOR>

A;Cross-references: EMBL:X70166; NID:964339; PID:964339
 C;Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match Score 44.5%; DB 2; Length 244;

Best Local Similarity 55.6%; Pred. No. 2.6%;
 Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 2 QYIKANS---KFIGITEL 16
 Db 51 EYIRFNSTVGKFVGYTEL 68.

RESULT 3
 S31029
 Gene 84 Protein - Mycobacterium phage L5

C;Species: Mycobacterium phage L5

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999

C;Accession: S31029
 R;Donnelly-Wu, M.K.; Jacobs Jr., W.R.; Hatfull, G.F.

Mol. Microbiol. 7, 407-417, 1993

A;Title: Superinfection immunity of mycobacteriophage L5: applications for genetic trans-

A;Reference number: S30949; PMID:93211283; MJD:8453767

A;Accession: S31029
 A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA
 A;Residues: 1-66 <DCN>

A;Cross-references: EMBL:Z18946; NID:915859; PID:CAA79460.1; PID:e59702; PID:9579152

RESULT 4
 G66826
 Query Match Score 44; DB 2; Length 66;

Best Local Similarity 57.1%; Pred. No. 0.06%;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YIKANSKFIGITEL 16
 Db 50 YIKRNGKFVGTWEV 63

RESULT 5
 T42976
 hypothetical protein 63 - ateline herpesvirus 3 (strain 73)

C;Species: ateline herpesvirus 3

C;Variety: strain 73

C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000

C;Accession: T42976
 R;Albrecht, J.C.; Fleckenstein, B.

Submitted to the EMBL Data Library, August 1998

A;Description: Primary structure of the herpesvirus atelies genome.

A;Accession number: 222274

A;Status: Preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: DNA
 A;Residues: 1-899 <ALB>

A;Cross-references: EMBL:AF003424; PID: AAC95587.1

A;Experimental source: strain 73

Query Match Score 43; DB 2; Length 899;

Best Local Similarity 64.3%; Pred. No. 18%;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QYIKANSKFIGITEL 15
 Db 124 QYITNSNTFVGTE 137

RESULT 6

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1992

C;Genetics:

A;Gene: 84
 A;Start codon: GTG

Query Match Score 44; DB 2; Length 66;

Best Local Similarity 57.1%; Pred. No. 0.06%;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YIKANSKFIGITEL 16
 Db 50 YIKRNGKFVGTWEV 63

RESULT 7
 G66826
 Query Match Score 44; DB 2; Length 66;

Best Local Similarity 57.1%; Pred. No. 0.06%;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YIKANSKFIGITEL 16
 Db 50 YIKRNGKFVGTWEV 63

RESULT 8
 G66826
 Query Match Score 43; DB 2; Length 180;

Best Local Similarity 69.2%; Pred. No. 3.6%;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 YIKANSKFIGITEL 16
 Db 65 IEANDDFIGIVEL 77

RESULT 9
 T42976
 hypothetical protein 63 - ateline herpesvirus 3 (strain 73)

C;Species: ateline herpesvirus 3

C;Variety: strain 73

C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000

C;Accession: T42976
 R;Albrecht, J.C.; Fleckenstein, B.

Submitted to the EMBL Data Library, August 1998

A;Description: Primary structure of the herpesvirus atelies genome.

A;Accession number: 222274

A;Status: Preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: DNA
 A;Residues: 1-899 <ALB>

A;Cross-references: EMBL:AF003424; PID: AAC95587.1

A;Experimental source: strain 73

Query Match Score 43; DB 2; Length 899;

Best Local Similarity 64.3%; Pred. No. 18%;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QYIKANSKFIGITEL 15
 Db 124 QYITNSNTFVGTE 137

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6
OM protein - protein search, using SW model
Run on: March 10, 2004, 08:58:54 ; Search time 4.98054 Seconds
Title: US-09-848-834A-2
Perfect score: 79
Sequence: 1 MQYIKANSKFIGITEL 16
Scoring table: BIOSUM62
GapOp 10.0 , GapExt 0.5
Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database : PIR 7B:
1: Pir1:
2: Pir2:
3: Pir3:
4: Pir4:
5: Pir5:
6: Pir6:
7: Pir7:
8: Pir8:
9: Pir9:
10: Pir10:
11: Pir11:
12: Pir12:
13: Pir13:
14: Pir14:
15: Pir15:
16: Pir16:
17: Pir17:
18: Pir18:
19: Pir19:
20: Pir20:
21: Pir21:
22: Pir22:
23: Pir23:
24: Pir24:
25: Pir25:
26: Pir26:
27: Pir27:
28: Pir28:
29: Pir29:

RESULT 1
BTCLTN tentoxilysin (EC 3.4.24.68) Precursor - Clostridium tetani
N,Alternative names: tetanus neurotoxin
C;Species: Clostridium tetani
C;Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 03-Jun-2002
C;Accession: A25689; S69349; S69344; S03364
R;Bisel, U.; Jarausch, W.; Goretzki, K.; Henschel, A.; Engels, J.; Weller, U.; Hudel, J.
EMBO J., 5, 2495-2502, 1986
A;Title: Tetanus toxin: primary structure, expression in E. coli, and homology with botulinum toxin
A;Reference number: A25757; MUID:8704047; PMID:1774547
A;Accession: A25689
A;Molecule type: DNA
A;Residues: 1-115 <EIS>
A;Cross-references: GB:X04436; NID:940769; PIDN:CAA28033.1; PID:940770
R;Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res., 14, 7809-7812, 1986
A;Title: The complete nucleotide sequence of tetanus toxin
A;Reference number: A25757; MUID:86085672; PMID:1510187
A;Accession: A25757
A;Molecule type: DNA
A;Residues: 1-1315 <PAI>
A;Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:940774
A;Experimental source: Strain CN3911
R;Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol., 165, 21-27, 1986
A;Title: Cloning, sequencing, and expression of tetanus toxin fragment C in MHC class II beta
A;Reference number: A25194; MUID:86053814; PMID:1536478
A;Accession: A25757
A;Molecule type: DNA
A;Residues: 743-1315 <FA2>
A;Cross-references: GB:M12739; NID:9144920; PIDN:AAA23282.1; PID:9144921
A;Accession: B25194
A;Molecule type: protein
A;Residues: 8-94 <PA3>
R;Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun., 57, 3588-3593, 1989
A;Title: Isolation, purification, and characterization of fragment B, the NH-2-terminal MHC class II beta
A;Reference number: A60759; MUID:90035436; PMID:2478476
A;Accession: A60759
A;Molecule type: protein
A;Residues: 461-475 <MAT>
R;Demote, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol., 142, 3940-402, 1989
A;Title: Deletion of several DR-restricted tetanus toxin T cell epitopes
A;Reference number: JS0098; MUID:89093918; PMID:2463305
A;Contents: annotation, epitope region
R;Schiavo, G.; Benfenati, F.; Poulat, B.; Laureto, O.; de Laureto, P.P.; DasGupta, B.
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by protein
A;Reference number: S27125; MUID:93063293; PMID:1331807
A;Contents: annotation

ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	1315	1 BTCLTN tentoxilysin (EC 3.4.24.68) Precursor - Clostridium tetani	
2	44.5	56.3	244	2 S29982 class II histocompatibility antigen 84 protein - gene 84 protein - diaminine N-acetyltransferase	
3	44	55.7	66	2 S31029	
4	43	54.4	180	2 G86826	
5	43	54.4	899	2 T42976	
6	42.5	53.8	82	2 I50966	
7	42.5	53.8	1060	2 S06286	
8	42.5	53.8	1086	2 S16752	
9	42.5	53.8	1701	2 A26868	
10	42.5	53.8	1701	2 A26868	
11	42.5	53.8	1726	1 SAZQGM	
12	42.5	53.8	1726	1 A45948	
13	42	53.2	1333	2 S38635	
14	41	51.9	79	2 D85794	
15	41	51.9	123	2 G48677	
16	41	51.9	447	2 H37146	
17	41	51.9	899	2 G36812	
18	40.5	51.3	245	2 G39980	
19	40	50.6	194	2 G64026	
20	40	50.6	213	1 K1YMC	
21	40	50.6	501	1 A96158	
22	40	50.6	501	2 TS2135	
23	40	50.6	601	1 A5485	
24	40	50.6	601	1 G68440	
25	40	50.6	644	2 S16746	
26	39	49.4	102	2 PH1491	
27	39	49.4	119	2 PH1516	
28	39	49.4	119	2 PH1518	
29	39	49.4	119	2 PH1519	

STRAIN=C57BL/6J; TISSUE=Head;
 RC Adachi J., Aikawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu M., Hiramoto K., Hirashita T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Iwasa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda N., Koya S., Kurihara C.,
 RA Matsuzaki T., Miyazaki R., Nishii K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shiragawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagaya A., Takayashi F., Tahaka T.,
 RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.,
 RA Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
 [2]

RN SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=23354633; PubMed=12466851;
 RA THE PHANTOM Consortium;
 RA RIKEN Genome Exploration Research Group Phase I & II Team:
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RT Nature 409:685-690 (2001).
 RL [4]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.,
 RA RIKEN FANTOM Consortium;
 RA "High-efficiency full-length cDNA cloning.";
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44 (1999).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
 RA "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630 (2000).
 [6]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itaya M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishii K., Kitazonai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Kazama M., Niizumi T., Harada A.,
 RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanuki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Muramatsu M., Hayashizaki Y.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RA "RIKEN integrated sequence analysis (RISA) system:384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 DR EMBL; AF014033; BAB29151.2; -.
 DR MGD; MGJ:1920407; Ced6.
 DR GO:0016491; F-oxidoreductase activity; IEA.
 DR GO:00008152; P-metabolism; IEA.
 DR InterPro; IPRO02086; Aldehyde_dehydr.
 DR InterPro; IPRO06030; PRB_PID.
 DR Pfam; PF00640; PID: 1.
 DR SMART; SM00462; PRB; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
 DR PROSITE; PS01179; PID: 1.
 DR SEQUENCE 291 AA; 32302 MW; 91FF3F41CD0BC6F2 CRC64;
 SQ Query Match 51.4%; Score 43; DB 11; Length 291;
 Best Local Similarity 57.1%; Pred. No. 25;

DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)	Qy 2 QYIKANS--KPIGIDTEL 16
DE MHC class II beta 1 (Fragment).	Db 33 EYIRFNSTVGKFVGYTEL 50
GN Salvelinus namaycush (lake trout).	
OS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;	
OC Protacanthopterygii; Neopterygii; Teleostei; Buteleosteii;	
OC Salmoniformes; Salmonidae; Salvelinus.	
NCBI_TaxID:8040; RN [1]	
RP SEQUENCE FROM N.A.	PRELIMINARY; PRT;
RA Dorschner M.O., Duris T., Phillips R.B.;	ID Q95HX1; ID Q95HX1;
RT "Diversity of a Lake Trout Mhc Class II Gens.";	AC Q95HX1;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.	DT 01-DEC-2001 (T-EMBLrel. 19, Created)
EMBL: AF13026; AAD08891; -.	DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DR GO; GO:0016021; C:integral to membrane; IEA.	DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DR GO; GO:0045012; F:MHC class II receptor activity; IEA.	DE MHC class II B antigen (Fragment).
DR GO; GO:0019884; P:antigen presentation, exogenous antigen via M. . . ; IEA.	GN DB1.
DR GO; GO:0019886; P:antigen processing, exogenous antigen via M. . . ; IEA.	OS Salmo salar (Atlantic Salmon).
DR GO; GO:0008555; P:immune response; IEA.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
DR Pfam: PF000353; MHC_II_beta.	OC Actinopterygii; Neopterygii; Teleostei; Buteleosteii;
DR ProDom; PF000369; MHC_II_beta; 1.	OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
DR Glycoprotein; MHC_II_Transmembrane.	NCBI_TaxID:8030; RN [1]
RW	RP SEQUENCE FROM N.A.
FT NON_TER 1 1	MEDLINE=21383619; PubMed=11491536;
FT NON_TER 71 71	RA Langefors A., Lohm J., von Schantz T.;
SEQUENCE 71 AA; 8002 MW; B3095286582A9F2D CRC64;	RT "Allelic polymorphism in MHC class II B in four populations of Atlantic salmon (Salmo salar).";
Query Match 56.3%; Score 44.5%; DB 7; Length 71;	RL Immunogenetics 53:329-336(2001);
Best Local Similarity 55.6%; Pred. No. 3.3;	DR ENB1; AF104370; AA040021; -.
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;	DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0045012; F:MHC class II receptor activity; IEA.	DR GO; GO:0045012; F:MHC class II receptor activity; IEA.
DR GO; GO:0019884; P:antigen presentation, exogenous antigen via M. . . ; IEA.	DR GO; GO:0019884; P:antigen presentation, exogenous antigen via M. . . ; IEA.
DR GO; GO:0019886; P:antigen processing, exogenous antigen via M. . . ; IEA.	DR GO; GO:0019886; P:antigen processing, exogenous antigen via M. . . ; IEA.
DR GO; GO:0008555; P:immune response; IEA.	DR InterPro; IPR000353; MHC_II_beta; 1.
DR Pfam: PF000353; MHC_II_beta; 1.	DR Pfam; PF000369; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.	DR ProdDom; PD000328; MHC_II_beta; 1.
RW	KW MHC.
FT NON_TER 1 1	
SEQUENCE 85 AA; 9729 MW; 27C977931F1F01C4 CRC64;	
Query Match 56.3%; Score 44.5%; DB 7; Length 85;	
Best Local Similarity 55.6%; Pred. No. 4;	
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;	
DR GO; GO:0016021; C:integral to membrane; IEA.	
DR GO; GO:0045012; F:MHC class II receptor activity; IEA.	
DR GO; GO:0019884; P:antigen presentation, exogenous antigen via M. . . ; IEA.	
DR GO; GO:0019886; P:antigen processing, exogenous antigen via M. . . ; IEA.	
DR GO; GO:0008555; P:immune response; IEA.	
DR InterPro; IPR000353; MHC_II_beta; 1.	
DR Pfam: PF000353; MHC_II_beta; 1.	
DR ProDom; PD000328; MHC_II_beta; 1.	
RW	RN [1]
RP SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.
RA Landry C., Bernatchez L.;	Q95HX4; PRELIMINARY; PRT;
RT "Comparative analysis of population structure across environments and geographic scales at Major Histocompatibility Complex and microsatellite in Atlantic salmon (Salmo salar).";	ID Q95HX4; ID Q95HX4;
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.	AC Q95HX4;
RL Salmo salar (Atlantic salmon).	DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DR GO; GO:0016021; C:integral to membrane; IEA.	DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DR GO; GO:0045012; F:MHC class II receptor activity; IEA.	DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DR GO; GO:0019884; P:antigen presentation, exogenous antigen via M. . . ; IEA.	DE MHC class II B antigen (Fragment).
DR GO; GO:0019886; P:antigen processing, exogenous antigen via M. . . ; IEA.	GN DB1.
DR GO; GO:0008555; P:immune response; IEA.	OS Salmo salar (Atlantic Salmon).
DR InterPro; IPR000353; MHC_II_beta; 1.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
DR Pfam: PF000353; MHC_II_beta; 1.	OC Actinopterygii; Neopterygii; Teleostei; Buteleosteii;
DR ProDom; PD000328; MHC_II_beta; 1.	OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
RW	NCBI_TaxID:8030; RN [1]
FT NON_TER 1 1	RP SEQUENCE FROM N.A.
FT NON_TER 85 85	MEDLINE=21383619; PubMed=11491536;
SEQUENCE 85 AA; 9743 MW; 3214E01AD1B66AC5 CRC64;	RA Langefors A., Lohm J., von Schantz T.;
Query Match 56.3%; Score 44.5%; DB 7; Length 85;	RT "Allelic polymorphism in MHC class II B in four populations of Atlantic salmon (Salmo salar).";
Best Local Similarity 55.6%; Pred. No. 4;	DR ENB1; AF104577; AA040091; -.
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;	DR GO; GO:0016021; C:integral to membrane; IEA.

Best Local Similarity	100.0%	Pred. No.	4.3e-05
Matches	16	Conservative	0;
Mismatches	0;	Indels	0;
Gaps	0;		
RESULT 2			
Q7YQH3	PRELIMINARY;	PRT;	447 AA.
Q7YQH3;			
01-OCT-2003 (TREMBLrel. 25, Created)			
01-OCT-2003 (TREMBLrel. 25, Last sequence update)			
01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
Endonuclease (EC 4.2.1.11).			
ENO OR BFL157.			
Candidatus Blochmannia floridanus.			
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
Enterobacteriaceae; ant endosymbionts; Candidatus Blochmannia.			
[1]			
SEQUENCE FROM N.A.			
Gil R., Silva P.J., Zientz E., Delmotte F., Gonzalez-Candela F.,			
Latorre A., Rausell C., Kamerbeek J., Gadu J., Hoildobler B.,			
van Ham R.C.H.J., Gross R., Moya A.,			
"The genome sequence of Blochmannia floridanus: comparative analysis of reduced genomes";			
Proc. Natl. Acad. Sci. U.S.A. 100:9338-9333 (2003).			
EMBL; BX248544; CAD83678.1; -.			
Lysine; Complete proteome.			
SEQUENCE 447 AA; 49005 MW; 465B69c3273C7AC4 CRC64;			
Query Match	58.2%	Score	46; DB 16; Length 447;
Best Local Similarity	46.7%	Pred. No.	11;
Matches	7;	Conservative	5; Mismatches
Indels	3;	Gaps	0;
Y	2 QYIKANSKFIGITEL 16		
416 EFLKDNSKFIGVNEI 430			
RESULT 3			
Q835J8	PRELIMINARY;	PRT;	880 AA.
Q835J8;			
01-JUN-2003 (TREMBLrel. 24, Created)			
01-OCT-2003 (TREMBLrel. 24, Last sequence update)			
Alanyl tRNA synthetase.			
ALAS OR EPI1379.			
Enterococcus faecalis (Streptococcaceae).			
Bacteria; Firmicutes; Lactobacillales; Enterococaceae; Enterococcus.			
[1]			
SEQUENCE FROM N.A.			
STRAIN V53 / ATCC 700802;			
MEDLINE=22550857; PubMed=12663927;			
Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,			
Read T.D., Roux D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,			
Tettelein H., Douson R.J., Umayam L., Brinkac L., Beanan M., Madupu R., Nelson W.,			
Daugherty S., Deboy R.T., Durkin S., Kolonay J., Utterback T., Hansen T., Shetty J., Khouri H.,			
Vamathevan J., Tran B., Upton J., Fraser C.M., Utterback T., Redline D., Ketchum K.A., Dougherty B.A., Fraser C.M.,			
"Role of mobile DNA in the evolution of vancomycin-resistant Enterococcus faecalis".			
Science 299:1207-12074 (2003).			
QL			
DR			
EMBL; AE01951; AA081170.1; -.			
TIGR; EF1379; -.			
DR			
GO; GO:0004813; P:alanine-tRNA ligase activity; IEA.			
DR			
GO; GO:0005224; P:cysteine-type endopeptidase activity; IEA.			
DP			
GO; GO:0004197; P:cysteine-type endopeptidase activity; IEA.			
RESULT 5			
Q8XRJ9	PRELIMINARY;	PRT;	71 AA.
ID			
Q8XRJ9;			
AC			
DR			
01-NOV-1999 (TREMBLrel. 12, Created)			
DT			
01-NOV-1999 (TREMBLrel. 12, Last sequence update)			

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 15.8132 Seconds

(without alignments)

319.45 Million cell updates/sec

Title: US-09-848-834A-2

Perfect score: 79

Sequence: 1 MOYIKANSKFIGITTEL 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:

1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_25:

```

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_minic:*
8: sp_orcanelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_reddent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*
```

```

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_minic:*
8: sp_orcanelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_reddent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*
```

RESULT 1

Q93N27

ID Q93N27

AC Q93N27;

DT 01-DEC-2001 (TRIMBLrel. 19, Created)

DT 01-DEC-2001 (TRIMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TRIMBLrel. 25, Last annotation update)

DR GO:0004816; F:endopeptidase inhibitor activity; IEA.

DR GO:0004827; F:metallopeptidase activity; IEA.

DR GO:0015101; F:toxin activity; IEA.

DR GO:0008220; F:zinc ion binding; IEA.

DR GO:0009405; P:protein binding; IEA.

DR InterPro: IPR001064; CrysFallin.

DR InterPro: IPR002160; Kunitz_league.

DR InterPro: IPR000395; Peptidase_M27.

DR InterPro: IPR006025; Pept_M_Zn_BS.

DR Pfam: PF01742; Peptidase_M27_1.

DR PRNTS; PRO0160; BONTONILISIN.

DR ProDom; PD000963; Bontoxilisin.

DR PROSITE; PS00225; CRYSTALLIN_BETA_GAMMA; 1.

DR PROSITE; PS01042; ZINC_PROTEASE; 1.

FT NON_TER 1

FT NON_TER 1310 1310

SEQUENCE 1310 AA; 150316 MW;

SQ 100.0% Score 79; DB 2; Length 1310;

SEQUENCE FROM N.A.
 RA Shumin Z., Dianliang L.;
 RT Cloning and sequence analysis of *Cetanus* toxin gene. ";
 RL Submitted (UN-2001) to the EMBL/GenBank/DBJ databases.
 DR AP389424; AAK72964.2; -
 DR GO:0004816; F:endopeptidase inhibitor activity; IEA.
 DR GO:0004827; F:metallopeptidase activity; IEA.
 DR GO:0015101; F:toxin activity; IEA.
 DR GO:0008220; F:zinc ion binding; IEA.
 DR GO:0009405; P:protein binding; IEA.
 DR InterPro: IPR001064; CrysFallin.
 DR InterPro: IPR002160; Kunitz_league.
 DR InterPro: IPR000395; Peptidase_M27.
 DR InterPro: IPR006025; Pept_M_Zn_BS.
 DR Pfam: PF01742; Peptidase_M27_1.
 DR PRNTS; PRO0160; BONTONILISIN.
 DR ProDom; PD000963; Bontoxilisin.
 DR PROSITE; PS00225; CRYSTALLIN_BETA_GAMMA; 1.
 DR PROSITE; PS01042; ZINC_PROTEASE; 1.
 FT NON_TER 1
 FT NON_TER 1310 1310
 SQ 100.0% Score 79; DB 2; Length 1310;

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	Q93N27	100.0	1310	2	Q93N27	Q93N27 clostridium
2	Q93N27	46	58.2	447	16	candidatus
3	Q93N27	46	58.2	890	16	O7VQH3
4	Q93N27	44.5	56.3	60	7	O31585
5	Q93N27	44.5	56.3	71	7	Q9XRJ9
6	Q93N27	44.5	56.3	85	7	Q951S2
7	Q93N27	44.5	56.3	85	7	Q95HY1
8	Q93N27	44.5	56.3	86	7	Q95HX4
9	Q93N27	44.5	56.3	244	7	Q31590
10	Q93N27	44	55.7	546	10	Q9XG37
11	Q93N27	43.5	55.1	67	7	Q31578
12	Q93N27	43	54.4	180	16	O9CF66
13	Q93N27	43	54.4	250	9	O9MCL7
14	Q93N27	43	54.4	252	11	Q9Xje8
15	Q93N27	43	54.4	291	11	lactococcus
16	Q93N27	43	54.4	304	11	Q9CrV4

```

; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
; NAME/KEY: PEPTIDE
; LOCATION: (32)..(37)
; OTHER INFORMATION: (Tetoxyllysine)
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(46)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
; US-09-848-834A-19
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(16)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
; Query Match 96.2%; Score 76; DB 9; Length 46;
; Best Local Similarity 93.8%; Pred. No. 1.1e-06;
; Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
; Qy 1 MOYIKANSKFIGITEL 16
; Db 16 IQYIKANSKFIGIGIBL 31
; RESULT 13
; US-10-297-942-8
; Sequence 8, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
; APPLICANT: Ferring BV
; TITLE OF INVENTION: Solubilised Protein Vaccines
; FILE REFERENCE: P68445US0
; CURRENT APPLICATION NUMBER: US/10/297,942
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: PCT/DK01/00431
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: DK PA 2000 00966
; PRIOR FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 8
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-297-942-8
; Query Match 96.2%; Score 76; DB 14; Length 158;
; Best Local Similarity 93.8%; Pred. No. 4.5e-06;
; Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
; Qy 1 MOYIKANSKFIGITEL 16
; Db 76 IQYIKANSKFIGITEL 91
; RESULT 14
; US-09-943-548-9
; Sequence 9, Application US/09943548
; Patent No. US20020042364A1
; GENERAL INFORMATION:
; APPLICANT: Ritterhaus, Charles W.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY
; FILE REFERENCE: TCS-411-1P US-1; TCS-411-1P US-2
; CURRENT APPLICATION NUMBER: US/09/943,548
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 09/943,334
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/943,548
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 9
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vaccine peptide of the invention
; US-10-339-522-9
; Query Match 94.9%; Score 75; DB 9; Length 50;
; Best Local Similarity 100.0%; Pred. No. 1.3e-06;
; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; Qy 1 MOYIKANSKFIGITEL 15
; Db 1 MOYIKANSKFIGITEL 15
; Search completed: March 10, 2004, 10:25:47
; Job time : 12.5759 secs

```

```

RESULT 9
US-10-451-596-141
; Sequence 141; Application US/10241596
; Publication No. US20030166238A1
; GENERAL INFORMATION:
; APPLICANT: Microbiological Research Authority
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581_0130003
; CURRENT APPLICATION NUMBER: US/10/241,596
; CURRENT FILING DATE: 2002-09-12
; PRIORITY APPLICATION NUMBER: US 09/255,829
; PRIORITY FILING DATE: 1999-02-23
; PRIORITY APPLICATION NUMBER: US 09/242,689
; PRIORITY FILING DATE: 1999-02-23
; PRIORITY APPLICATION NUMBER: PCT/GB97/02273
; PRIORITY FILING DATE: 1997-08-22
; PRIORITY APPLICATION NUMBER: US 08/782,893
; PRIORITY FILING DATE: 1996-12-27
; PRIORITY APPLICATION NUMBER: GB 9625996 5
; PRIORITY FILING DATE: 1996-12-13
; PRIORITY APPLICATION NUMBER: GB 9617671 4
; PRIORITY FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 175
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 141
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Clostridium tetani
; US-10-241-596-141

Query Match Score 79; DB 14; Length 1315;
Best Local Similarity 100%; Pred. No. 1.4e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQYIKANSKFIGITEL 16
Db 829 MQYIKANSKFIGITEL 844

RESULT 10
US-10-452-024-145
; Sequence 145; Application US/10452024
; Publication No. US20040013687A1
; GENERAL INFORMATION:
; APPLICANT: Simpson, Lance
; APPLICANT: Park, Jung-Beak
; APPLICANT: Maksynowich, Andrew
; TITLE OF INVENTION: Compositions and Methods For Transepithelial Molecular Transport
; FILE REFERENCE: 9B55-960U
; CURRENT APPLICATION NUMBER: US/10/452,024
; CURRENT FILING DATE: 2003-05-02
; PRIORITY APPLICATION NUMBER: 60/384,949
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 188
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 145
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Clostridium tetani
; US-10-452-024-145

Query Match Score 79; DB 15; Length 1315;
Best Local Similarity 100%; Pred. No. 1.4e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQYIKANSKFIGITEL 16
Db 829 MQYIKANSKFIGITEL 844

RESULT 11
US-10-452-024-146
; Sequence 146; Application US/10452024
; Publication No. US20040013687A1
; GENERAL INFORMATION:
; APPLICANT: Simpson, Lance
; APPLICANT: Park, Jung-Beak
; APPLICANT: Maksynowich, Andrew
; TITLE OF INVENTION: Compositions and Methods For Transepithelial Molecular Transport
; FILE REFERENCE: 9B55-960U
; CURRENT APPLICATION NUMBER: US/10/452,024
; CURRENT FILING DATE: 2003-05-02
; PRIORITY APPLICATION NUMBER: 60/384,949
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 188
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 146
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Clostridium tetani
; US-10-452-024-146

Query Match Score 79; DB 15; Length 1315;
Best Local Similarity 100%; Pred. No. 1.4e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQYIKANSKFIGITEL 16
Db 829 MQYIKANSKFIGITEL 844

RESULT 12
US-09-848-834A-19
; Sequence 19; Application US/09848834A
; Patent No. US200200764-6A1
; GENERAL INFORMATION:
; APPLICANT: Abtton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102855-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIORITY APPLICATION NUMBER: 60/205,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 19
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human GnRH linked by a spacer to amino acid sequence 830-844 of Tenteroxylsin
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tenteroxylsin (Tenteroxylsin) linked by a spacer to amino acid sequence 1-10 of human GnRH
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD-RES
; LOCATION: (1)-(1)
; OTHER INFORMATION: Amidated glycine or glycaminide
; NAME/KEY: MOD-RES
; LOCATION: (46)-(46)
; OTHER INFORMATION: Amidated glycine or glycaminide

Qy 1 MQYIKANSKFIGITEL 16
Db 16 IQLYIKANSKFIGITEL 31

RESULT 13
US-09-848-834A-19
; Sequence 19; Application US/09848834A
; Patent No. US200200764-6A1
; GENERAL INFORMATION:
; APPLICANT: Abtton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102855-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIORITY APPLICATION NUMBER: 60/205,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 19
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human GnRH linked by a spacer to amino acid sequence 830-844 of Tenteroxylsin
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tenteroxylsin (Tenteroxylsin) linked by a spacer to amino acid sequence 1-10 of human GnRH
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD-RES
; LOCATION: (1)-(1)
; OTHER INFORMATION: Amidated glycine or glycaminide
; NAME/KEY: MOD-RES
; LOCATION: (46)-(46)
; OTHER INFORMATION: Amidated glycine or glycaminide

Qy 1 MQYIKANSKFIGITEL 16
Db 16 IQLYIKANSKFIGITEL 31

```

APPLICANT: Microbiological Research Authority
 APPLICANT: The Sperrywood Laboratory Limited
 TITLE OF INVENTION: Recombinant Toxin Fragments
 FILE REFERENCE: 1581-0130003
 CURRENT APPLICATION NUMBER: US/10/241,596
 PRIOR APPLICATION NUMBER: US 09/255,829
 PRIOR FILING DATE: 1999-02-23
 PRIOR APPLICATION NUMBER: US 09/242,689
 PRIOR FILING DATE: 1999-02-23
 PRIOR APPLICATION NUMBER: PCT/GB97/022273
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: US 08/782,893
 PRIOR FILING DATE: 1996-12-27
 PRIOR APPLICATION NUMBER: GB 9625996.5
 PRIOR FILING DATE: 1996-12-13
 PRIOR APPLICATION NUMBER: GB 9617671.4
 PRIOR FILING DATE: 1996-08-23
 NUMBER OF SEQ ID NOS: 175
 SOFTWARE: Patentin version 3.1
 SEQ ID NO: 145
 LENGTH: 872
 TYPE: PRT
 ORGANISM: Clostridium tetani
 US-10-241-596-145

Query Match 100.0%; Score 79; DB 14; Length 872;
 Best Local Similarity 100.0%; Pred. No. 8.7e-06; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0;

Qy 1 MOYIKANSKFIGITEL 16
 Db 822 MOYIKANSKFIGITEL 837

RESULT 6
 US-10-241-596-143
 Sequence 143 Application US/10241596
 Publication No. US20030166238A1

GENERAL INFORMATION:
 APPLICANT: Microbiological Research Authority
 TITLE OF INVENTION: Recombinant Toxin Fragments
 FILE REFERENCE: 1581-0130003
 CURRENT APPLICATION NUMBER: US/10/241,596
 PRIOR FILING DATE: 2002-09-12
 PRIOR APPLICATION NUMBER: US 09/255,829
 PRIOR FILING DATE: 1999-02-23
 PRIOR APPLICATION NUMBER: US 09/242,689
 PRIOR FILING DATE: 1999-02-23
 PRIOR APPLICATION NUMBER: PCT/GB97/022273
 PRIOR FILING DATE: 1997-08-22
 PRIOR APPLICATION NUMBER: GB 9625996.5
 PRIOR FILING DATE: 1996-12-13
 NUMBER OF SEQ ID NOS: 175
 SOFTWARE: Patentin version 3.1
 SEQ ID NO: 143
 LENGTH: 879
 TYPE: PRT
 ORGANISM: Clostridium tetani
 US-10-241-596-143

Query Match 100.0%; Score 79; DB 15; Length 1310;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0;

Qy 1 MOYIKANSKFIGITEL 16
 Db 829 MOYIKANSKFIGITEL 844

GENERAL INFORMATION:
 APPLICANT: Hickey, William F.
 Griffin, Ann C.
 TITLE OF INVENTION: Proinsulin Peptide Compounds for Detecting and Treating Type I Diabetes
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street, Suite 510
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/346,563
 FILING DATE: 16-Jan-2003
 CLASSIFICATION: 424
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/08/472,704
 FILING DATE: 06-Jun-1995
 REGISTRATION NUMBER: 31,503
 APPLICATION NUMBER: US/08/272,220
 FILING DATE: 08-JULY-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: DeConti, Giulio A., Jr.
 REGISTRATION NUMBER: 31,503
 REFERENCE/DOCKET NUMBER: DCI-092DV
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEX/FAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 17 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: Peptide
 FRAGMENT TYPE: Internal
 FRAGMENT DESCRIPTION: SEQ ID NO: 23:
 US-10-321-717-23

Query Match Score 79; DB 15; Length 17;
 Best Local Similarity 100.0%; Pred. No. 1e-7;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MOYIKANSKFIGITEL 16
 Db 2 MQYIKANSKFIGITEL 17

RESULT 4
 US-10-452-024-177
 Sequence 177, Application US/10452024
 Publication No. US2004013687A1
 GENERAL INFORMATION:
 APPLICANT: Simpson, Lance
 APPLICANT: Park, Jung-Beak
 APPLICANT: Maksmowich, Andrew
 TITLE OF INVENTION: Compositions and Methods For Transepithelial Molecular Transport
 FILE REFERENCE: 9855-36U1
 CURRENT APPLICATION NUMBER: US/10/452,024
 CURRENT FILING DATE: 2003-06-02
 PRIOR APPLICATION NUMBER: 60/384,949
 PRIOR FILING DATE: 2002-05-31
 NUMBER OF SEQ ID NOS: 188
 SOFTWARE: Patentin version 3.2
 SEQ ID NO: 177
 LENGTH: 573
 TYPE: PPT
 ORGANISM: Clostridium tetani
 US-10-452-024-177

Query Match Score 79; DB 15; Length 573;
 Best Local Similarity 100.0%; Pred. No. 5.4e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MOYIKANSKFIGITEL 16
 Db 87 MQYIKANSKFIGITEL 102

RESULT 5
 US-10-241-596-145
 Sequence 145, Application US/10241596
 Publication No. US20030166238A1
 GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 10, 2004, 09:16:59 ; Search time 12.5759 Seconds

(without alignments)
268.645 Million cell updates/sec

Perfect score: 79 US-09-848-834a-2

Sequence: 1 MQYIKANSKPIGTEL 16

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : Published Applications AAI,* 1: /cn2_6_ptodata/2/pubpaas/US07_PUBCOMB.pep:*

2: /cn2_6_ptodata/2/pubpaas/PTCTUS_PUB.pep:*

3: /cn2_6_ptodata/2/pubpaas/US06_PUB.pep:*

4: /cn2_6_ptodata/2/pubpaas/US07_PUBCOMB.pep:*

5: /cn2_6_ptodata/2/pubpaas/US07_PUBCOMB.pep:*

6: /cn2_6_ptodata/2/pubpaas/PTCTUS_PUBCOMB.pep:*

7: /cn2_6_ptodata/2/pubpaas/US08_PUB.pep:*

8: /cn2_6_ptodata/2/pubpaas/US08_PUBCOMB.pep:*

9: /cn2_6_ptodata/2/pubpaas/US09_PUBCOMB.pep:*

10: /cn2_6_ptodata/2/pubpaas/US09E_PUBCOMB.pep:*

11: /cn2_6_ptodata/2/pubpaas/US09C_PUBCOMB.pep:*

12: /cn2_6_ptodata/2/pubpaas/US09_NEW_PUB.pep:*

13: /cn2_6_ptodata/2/pubpaas/US10A_PUBCOMB.pep:*

14: /cn2_6_ptodata/2/pubpaas/US10B_PUBCOMB.pep:*

15: /cn2_6_ptodata/2/pubpaas/US10C_PUBCOMB.pep:*

16: /cn2_6_ptodata/2/pubpaas/US10_NEW_PUB.pep:*

17: /cn2_6_ptodata/2/pubpaas/US60_NEW_PUB.pep:*

18: /cn2_6_ptodata/2/pubpaas/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	%
1	79	100.0	16	9 US-09-848-834a-2	Sequence 2, Appli	
2	79	100.0	17	15 US-10-346-563-23	Sequence 23, Appli	
3	79	100.0	17	15 US-10-321-717-23	Sequence 23, Appli	
4	79	100.0	573	15 US-10-452-024-177	Sequence 177, Appli	
5	79	100.0	872	14 US-10-241-596-145	Sequence 145, Appli	
6	79	100.0	879	14 US-10-241-596-143	Sequence 143, Appli	
7	79	100.0	887	14 US-10-241-596-147	Sequence 147, Appli	
8	79	100.0	1310	15 US-10-452-024-149	Sequence 149, Appli	
9	79	100.0	1315	15 US-10-241-596-141	Sequence 141, Appli	
10	79	100.0	1315	15 US-10-452-024-145	Sequence 145, Appli	
11	76	96.2	31	9 US-09-848-834a-15	Sequence 15, Appli	
12	76	96.2	46	9 US-09-848-834a-19	Sequence 19, Appli	
13	76	96.2	158	14 US-10-297-942-8	Sequence 8, Appli	
14	75	94.9	50	9 US-09-943-548-9	Sequence 9, Appli	
15	75	94.9	50	14 US-10-339-522-9	Sequence 9, Appli	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	%
1	79	100.0	16	9 US-09-848-834a-2	Sequence 2, Appli	
2	79	100.0	17	15 US-10-346-563-23	Sequence 23, Appli	
3	79	100.0	17	15 US-10-321-717-23	Sequence 23, Appli	
4	79	100.0	573	15 US-10-452-024-177	Sequence 177, Appli	
5	79	100.0	872	14 US-10-241-596-145	Sequence 145, Appli	
6	79	100.0	879	14 US-10-241-596-143	Sequence 143, Appli	
7	79	100.0	887	14 US-10-241-596-147	Sequence 147, Appli	
8	79	100.0	1310	15 US-10-452-024-149	Sequence 149, Appli	
9	79	100.0	1315	15 US-10-241-596-141	Sequence 141, Appli	
10	79	100.0	1315	15 US-10-452-024-145	Sequence 145, Appli	
11	76	96.2	31	9 US-09-848-834a-15	Sequence 15, Appli	
12	76	96.2	46	9 US-09-848-834a-19	Sequence 19, Appli	
13	76	96.2	158	14 US-10-297-942-8	Sequence 8, Appli	
14	75	94.9	50	9 US-09-943-548-9	Sequence 9, Appli	
15	75	94.9	50	14 US-10-339-522-9	Sequence 9, Appli	

Result No.	Score	Query Match	Length	DB ID	Description	%
1	79	100.0	16	9 US-09-848-834a-2	Sequence 2, Appli	
2	79	100.0	17	15 US-10-346-563-23	Sequence 23, Appli	
3	79	100.0	17	15 US-10-321-717-23	Sequence 23, Appli	
4	79	100.0	573	15 US-10-452-024-177	Sequence 177, Appli	
5	79	100.0	872	14 US-10-241-596-145	Sequence 145, Appli	
6	79	100.0	879	14 US-10-241-596-143	Sequence 143, Appli	
7	79	100.0	887	14 US-10-241-596-147	Sequence 147, Appli	
8	79	100.0	1310	15 US-10-452-024-149	Sequence 149, Appli	
9	79	100.0	1315	15 US-10-241-596-141	Sequence 141, Appli	
10	79	100.0	1315	15 US-10-452-024-145	Sequence 145, Appli	
11	76	96.2	31	9 US-09-848-834a-15	Sequence 15, Appli	
12	76	96.2	46	9 US-09-848-834a-19	Sequence 19, Appli	
13	76	96.2	158	14 US-10-297-942-8	Sequence 8, Appli	
14	75	94.9	50	9 US-09-943-548-9	Sequence 9, Appli	
15	75	94.9	50	14 US-10-339-522-9	Sequence 9, Appli	

Result No.	Score	Query Match	Length	DB ID	Description	%
1	79	100.0	16	9 US-09-848-834a-2	Sequence 2, Appli	
2	79	100.0	17	15 US-10-346-563-23	Sequence 23, Appli	
3	79	100.0	17	15 US-10-321-717-23	Sequence 23, Appli	
4	79	100.0	573	15 US-10-452-024-177	Sequence 177, Appli	
5	79	100.0	872	14 US-10-241-596-145	Sequence 145, Appli	
6	79	100.0	879	14 US-10-241-596-143	Sequence 143, Appli	
7	79	100.0	887	14 US-10-241-596-147	Sequence 147, Appli	
8	79	100.0	1310	15 US-10-452-024-149	Sequence 149, Appli	
9	79	100.0	1315	15 US-10-241-596-141	Sequence 141, Appli	
10	79	100.0	1315	15 US-10-452-024-145	Sequence 145, Appli	
11	76	96.2	31	9 US-09-848-834a-15	Sequence 15, Appli	
12	76	96.2	46	9 US-09-848-834a-19	Sequence 19, Appli	
13	76	96.2	158	14 US-10-297-942-8	Sequence 8, Appli	
14	75	94.9	50	9 US-09-943-548-9	Sequence 9, Appli	
15	75	94.9	50	14 US-10-339-522-9	Sequence 9, Appli	

Result No.	Score	Query Match	Length	DB ID	Description	%
1	79	100.0	16	9 US-09-848-834a-2	Sequence 2, Appli	
2	79	100.0	17	15 US-10-346-563-23	Sequence 23, Appli	
3	79	100.0	17	15 US-10-321-717-23	Sequence 23, Appli	
4	79	100.0	573	15 US-10-452-024-177	Sequence 177, Appli	
5	79	100.0	872	14 US-10-241-596-145	Sequence 145, Appli	
6	79	100.0	879	14 US-10-241-596-143	Sequence 143, Appli	
7	79	100.0	887	14 US-10-241-596-147	Sequence 147, Appli	
8	79	100.0	1310	15 US-10-452-024-149	Sequence 149, Appli	
9	79	100.0	1315	15 US-10-241-596-141	Sequence 141, Appli	
10	79	100.0	1315	15 US-10-452-024-145	Sequence 145, Appli	
11	76	96.2	31	9 US-09-848-834a-15	Sequence 15, Appli	
12	76	96.2	46	9 US-09-848-834a-19	Sequence 19, Appli	
13	76	96.2	158	14 US-10-297-942-8	Sequence 8, Appli	
14	75	94.9	50	9 US-09-943-548-9	Sequence 9, Appli	
15	75	94.9	50	14 US-10-339-522-9	Sequence 9, Appli	

Result No.	Score	Query Match	Length	DB ID	Description	%
1	79	100.0	16	9 US-09-848-834a-2	Sequence 2, Appli	
2	79	100.0	17	15 US-10-346-563-23	Sequence 23, Appli	
3	79	100.0	17	15 US-10-321-717-23	Sequence 23, Appli	
4	79	100.0	573	15 US-10-452-024-177	Sequence 177, Appli	
5	79	100.0	872	14 US-10-241-596-145	Sequence 145, Appli	
6	79	100.0	879	14 US-10-241-596-143	Sequence 143, Appli	
7	79	100.0	887	14 US-10-241-596-147	Sequence 147, Appli	
8	79	100.0	1310	15 US-10-452-024-149	Sequence 149, Appli	
9	79	100.0	1315	15 US-10-241-596-141	Sequence 141, Appli	
10	79	100.0	1315	15 US-10-452-024-145	Sequence 145, Appli	
11	76	96.2	31	9 US-09-848-834a-15	Sequence 15, Appli	
12	76	96.2	46	9 US-09-848-834a-19	Sequence 19, Appli	
13	76	96.2	158	14 US-10-297-942-8	Sequence 8, Appli	
14	75	94.9	50	9 US-09-943-548-9	Sequence 9, Appli	
15	75	94.9	50	14 US-10-339-522-9	Sequence 9, Appli	

Result No.	Score	Query Match	Length	DB ID	Description	%
1	79	100.0	16	9 US-09-848-834a-2	Sequence 2, Appli	
2	79	100.0	17	15 US-10-346-563-23	Sequence 23, Appli	
3	79	100.0	17	15 US-10-321-717-23	Sequence 23, Appli	
4	79	100.0	573	15 US-10-452-024-177	Sequence 177, Appli	
5	79	100.0	872	14 US-10-241-596-145	Sequence 145, Appli	
6	79	100.0	879	14 US-10-241-596-143	Sequence 143, Appli	
7	79	100.0	887	14 US-10-241-596-147	Sequence 147, Appli	
8	79	100.0	1310	15 US-10-452-024-149	Sequence 149, Appli	
9	79	100.0	1315	15 US-10-241-596-141	Sequence 141, Appli	
10						

DR GO:0003793; P: defense/immunity protein activity; IEA.
 DR InterPro; IPR003067; Circmsprzoite.
 DR STRAIN=FUP/SP; and MAD01;
 RC de Stricker K.; Vuast J.; Jepsen S.; Oeuvray C.; Theisen M.;
 RA "Sequence variation in the non-repeat region of the Plasmodium
 RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
 RT Burmese field isolates and from laboratory strains."
 RT Submitted (SBP-1999) to the EMBL/GenBank/DDBJ databases.
 RL EMBL; AJ26953; CAB64178.1; -.
 DR EMBL; AJ26953; CAB64175.1; -.
 GO; GO:0003793; F: defense/immunity protein activity; IEA.
 DR InterPro; IPR003067; Circmsprzoite.
 DR InterPro; IPR000884; TSPI1.
 DR InterPro; IPR000884; TSPI1.

Query Match 85.3%; Score 81; DB 5; Length 80;
 Best Local Similarity 89.5%; Pred. No. 3.5e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0; Gaps 0;
 CS.

Qy 2 EKKIAKMEKASSVNVNS 20
 Db 62 EKKICKMERCSSSVNVNS 80

RESULT 13
 Q9U0Q4 PRELIMINARY; PRT; 80 AA.
 ID Q9U0Q4; PRELIMINARY; PRT; 80 AA.
 AC Q9U0Q4; PRELIMINARY; PRT; 80 AA.
 DT 01-MAY-2000 (TREMBLrel. 13; Created)
 DT 01-MAY-2000 (TREMBLrel. 13; Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24; Last annotation update)
 DE Circumsporozoite protein (fragment).
 GN CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_TaxID=5833;
 RN [1]_SEQUENCE FROM N.A.
 RP STRAIN=BL;
 RC de Stricker K.; Vuast J.; Jepsen S.; Oeuvray C.; Theisen M.;
 RA "Sequence variation in the non-repeat region of the Plasmodium
 RT falciparum Glutamate rich protein (GLURP) from Brazil, Senegalese, and
 RT Burmese field isolates and from laboratory strains."
 RT Submitted (SBP-1999) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AJ26941; CAB64167.1; -.
 GO; GO:0003793; F: defense/immunity protein activity; IEA.
 DR InterPro; IPR003067; Circmsprzoite.
 DR InterPro; IPR000884; TSPI1.
 DR STRAIN=FUP/SP; and MAD01;
 RC de Stricker K.; Vuast J.; Jepsen S.; Oeuvray C.; Theisen M.;
 RA "Sequence variation in the non-repeat region of the Plasmodium
 RT falciparum Glutamate rich protein (GLURP) from Brazil, Senegalese, and
 RT Burmese field isolates and from laboratory strains."
 RT Submitted (SBP-1999) to the EMBL/GenBank/DDBJ databases.
 RL EMBL; AJ26963; CAB64182.1; -.
 DR GO; GO:0003793; F: defense/immunity protein activity; IEA.
 DR InterPro; IPR003067; Circmsprzoite.
 DR InterPro; IPR000884; TSPI1.
 DR PRINTS; PR01303; CircMSPRZOITE.
 DR SMART; SM00203; TSPI1.1.
 FT NON-TER 1 1
 FT NON-TER 80 80
 SQ SEQUENCE 80 AA; 9102 MW; A3283B70CBE50FDE CRC64;

Query Match 85.3%; Score 81; DB 5; Length 80;
 Best Local Similarity 89.5%; Pred. No. 3.5e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0; Gaps 0;
 CS.

Qy 2 EKKIAKMEKASSVNVNS 20
 Db 62 EKKICKMERCSSSVNVNS 80

RESULT 14
 Q9TW00 PRELIMINARY; PRT; 80 AA.
 ID Q9TW00; PRELIMINARY; PRT; 80 AA.
 AC Q9TW00; PRELIMINARY; PRT; 80 AA.
 DT 01-MAY-2000 (TREMBLrel. 13; Created)
 DT 01-MAY-2000 (TREMBLrel. 13; Last sequence update)
 DE Circumsporozoite protein (fragment).
 GN CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_TaxID=5833;

Search completed: March 10, 2004, 09:25:23
 Job time : 20.7665 secs

RESULT 10
 Q9TVB9; PRELIMINARY; PRT; 80 AA.
 AC DT 01-MAY-2000 (TREMBUREL. 13, Created)
 RC DT 01-MAY-2000 (TREMBUREL. 13, Last sequence update)
 RA DT 01-JUN-2003 (TREMBUREL. 24, Last annotation update)
 RT Circumsporozoite protein (Fragment).
 CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_TAXID=5833;
 RN [1]
 SEQUENCE FROM N.A.
 STRAIN=M0, M6, M7, and M1;
 RC de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RA "Sequence variation in the non-repeat region of the Plasmodium falciparum Glutamate rich protein (GLURP) from Brazil, Senegalese, and Burmese field isolates and from laboratory strains.";
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RL CAB64242;
 DR EMBL; AJ269955;
 DR EMBL; AJ269958;
 DR EMBL; AJ269959;
 DR EMBL; AJ269956;
 DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
 DR InterPro; IPR003067; Circmsprzoite.
 DR PRIMs; PR01303; CRMSPRZOITE.
 DR Pfam; PF00090; tsp_1.
 DR SMART; SM0209; TSP1_1.
 DR PROSITE; PS50032; TSP1_1.
 FT NON_TER 1
 PT NON_TER 80
 SQ SEQUENCE 80 AA; 9074 MW; A8F40C90DB1C033E CRC64;
 Query Match 85.3%; Score 81; DB 5; Length 80;
 Best Local Similarity 89.5%; Pred. No. 3.5e-05;
 Matches 17; Conservative 0; Nismatches 2; Indels 0; Gaps 0;
 CS.

Q9U000; PRELIMINARY; PRT; 80 AA.
 AC DT 01-MAY-2000 (TREMBUREL. 13, Created)
 RC DT 01-MAY-2000 (TREMBUREL. 13, Last sequence update)
 RA DT 01-JUN-2003 (TREMBUREL. 24, Last annotation update)
 RT Circumsporozoite protein (Fragment).
 CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_TAXID=5833;
 RN [1]
 SEQUENCE FROM N.A.
 STRAIN=D28;
 RC de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RA "Sequence variation in the non-repeat region of the Plasmodium falciparum Glutamate rich protein (GLURP) from Brazil, Senegalese, and Burmese field isolates and from laboratory strains.";
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RL CAB6173_1;
 DR EMBL; AJ269948;
 DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
 DR InterPro; IPR003067; Circmsprzoite.
 DR PRIMs; PR01303; CRMSPRZOITE.
 DR SMART; SM0009; TSP1_1.
 DR PROSITE; PS50032; TSP1_1.
 FT NON_TER 1
 PT NON_TER 80
 SQ SEQUENCE 80 AA; 9000 MW; 03799BD18F0BF3BB3 CRC64;

RESULT 11
 Q9U005; PRELIMINARY; PRT; 80 AA.
 AC DT 01-MAY-2000 (TREMBUREL. 13, Created)
 RC DT 01-MAY-2000 (TREMBUREL. 13, Last sequence update)
 RA DT 01-JUN-2003 (TREMBUREL. 24, Last annotation update)
 RT Circumsporozoite protein (Fragment).
 CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_TAXID=5833;
 RN [1]
 SEQUENCE FROM N.A.
 STRAIN=D28;
 RC de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RA "Sequence variation in the non-repeat region of the Plasmodium falciparum Glutamate rich protein (GLURP) from Brazil, Senegalese, and Burmese field isolates and from laboratory strains.";
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RL CAB6173_1;
 DR EMBL; AJ269948;
 DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
 DR InterPro; IPR003067; Circmsprzoite.
 DR PRIMs; PR01303; CRMSPRZOITE.
 DR SMART; SM0009; TSP1_1.
 DR PROSITE; PS50032; TSP1_1.
 FT NON_TER 1
 PT NON_TER 80
 SQ SEQUENCE 80 AA; 9000 MW; 03799BD18F0BF3BB3 CRC64;

RESULT 12
 Q9TW76; PRELIMINARY; PRT; 80 AA.
 AC DT 01-MAY-2000 (TREMBUREL. 13, Created)
 RC DT 01-MAY-2000 (TREMBUREL. 13, Last sequence update)
 RA DT 01-JUN-2003 (TREMBUREL. 24, Last annotation update)
 RT Circumsporozoite protein (Fragment).
 CS.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_TAXID=5833;
 RN [1]
 SEQUENCE FROM N.A.
 STRAIN=PALO ALTO, and D50;
 RC de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
 RA "Sequence variation in the non-repeat region of the Plasmodium falciparum Glutamate rich protein (GLURP) from Brazil, Senegalese, and Burmese field isolates and from laboratory strains.";
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RL CAB44179_1;
 DR EMBL; AJ269944; CAB6170_1;
 DR EMBL; AJ269944; CAB6170_1; -.

DR GO; GO:0003793; P: defense/immunity protein activity; IEA.

DR InterPro; IPR003667; Circmsporzoite.

DR InterPro; IPR005884; TSP1.

DR PRINTS; PF00050; CSP1; 1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00209; TSP1; 1.

DR SMART; SM00092; TSP1; 1.

FT NON-TER 1 1

FT NON-TER 80 80

SEQ SEQUENCE 80 AA; 9000 MW; 036C7E119AFFB13B3 CRC64;

Query Match Score 81; DB 5; Length 80;
Best Local Similarity 89.5%; Pred. No. 3.5e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EKKIAKMEKASSVFNVVNS 20

Db 62 EKKICKMEKCSSVFNVVNS 80

RESULT 6 Q9UOP1 PRELIMINARY; PRT; 80 AA.

ID Q9UOP1; AC Q9UOP1; DT 01-MAY-2000 (TREMBrel. 13, Created)

DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)

DT 01-JUN-2003 (TREMBrel. 24, Last annotation update)

DE Circumsporozoite protein (Fragment).

CS

OS Plasmodium falciparum.

OC Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;

RN SEQUENCE FROM N.A.

RP STRAIN=N.A.

RC MEDLINE=2054; PubMed=11087922;

RA de Stricker K., Vuist J., Jepsen S., Oeuuvray C., Theisen M.;

"Conservation and heterogeneity of the glutamate-rich protein (GLURP)

among field isolates and laboratory lines of Plasmodium falciparum."

Mol. Biochem. Parasit. 111:123-130(2000).

DR EMBL; AU2269978; CAB64197; 1.

DR GO; GO:0003793; P: defense/immunity protein activity; IEA.

DR InterPro; IPR003067; Circmsporzoite.

DR InterPro; IPR00090; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00209; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

DR PROSITE; PS50092; TSP1.

DR PRINTS; PR01303; CRCMSPRZOTE.

DR SMART; SM00092; TSP1.

RESULT 2

Q910P3	PRELIMINARY;	PRT;	80 AA.
ID			
AC			
DT	01-MAY-2000 (TREMBrel. 13; Created)		
DT	01-MAY-2000 (TREMBrel. 13; Last sequence update)		
DT	01-JUN-2003 (TREMBrel. 24; Last annotation update)		
DE	Circumsporozoite protein (Fragment).		
GN	OS		
OC	Plasmodium falciparum		
OX	Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
RN	NCBI_TaxID=5833;		
RP	SEQUENCE FROM N.A.		
RC	STRAN=D4405;		
RA	de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.; "Sequence variation in the non-repeat region of the Plasmodium falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and Burmese field isolates and from laboratory strains"; Submitted (SBP-1999) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AJ269969; CAB64188.1; -.		
DR	GO; GO:0003793; F:defense/immunity protein activity; IEA.		
DR	InterPro; IPR003067; Crampsprotoite.		
DR	InterPro; IPR00884; TSP1.		
DR	Pfam; PF00090; tsp_1; 1.		
DR	PRINTS; PRO1303; CRCMSPRZOITE.		
DR	SMART; SM00209; TSP1; 1.		
DR	PROSITE; PS50032; TSP1; 1.		
FT	NON_TER 80 80		
SQ	SEQUENCE 80 AA; 9033 MW; BA71EBE0DB03193B CRC64;		

Query Match 85.3%; Score 81; DB 5; Length 80;
 Best Local Similarity 89.5%; Pred. No. 3.5e-05; Indels 0; Gaps 0;
 Matches 17; Conservative 0; MisMatches 2; Indels 0; Gaps 0;

Qy	2 EKKIARMEKASSVTVNVS 20		
Db	62 EKKIKRMKEKCSVTVNVS 80		

RESULT 3

Q9TW01	PRELIMINARY;	PRT;	80 AA.
ID			
AC			
DT	01-MAY-2000 (TREMBrel. 13; Created)		
DT	01-MAY-2000 (TREMBrel. 13; Last sequence update)		
DT	01-JUN-2003 (TREMBrel. 24; Last annotation update)		
DE	Circumsporozoite protein (Fragment).		
GN	OS		
OC	Plasmodium falciparum		
OX	Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
RN	NCBI_TaxID=5833;		
RP	SEQUENCE FROM N.A.		
RC	STRAN=M11; and D7;		
RA	de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.; "Sequence variation in the non-repeat region of the Plasmodium falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and Burmese field isolates and from laboratory strains"; Submitted (SBP-1999) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AJ269969; CAB64172.1; -.		
DR	GO; GO:0003793; F:defense/immunity protein activity; IEA.		
DR	InterPro; IPR003067; Crampsprotoite.		
DR	InterPro; IPR000884; TSP1.		
DR	Pfam; PF00090; tsp_1; 1.		
DR	PRINTS; PRO1303; CRCMSPRZOITE.		
DR	SMART; SM00209; TSP1; 1.		
DR	PROSITE; PS50032; TSP1; 1.		
FT	NON_TER 1 80		
SQ	SEQUENCE 80 AA; 9047 MW; BA7639290DB031C3B CRC64;		

Query Match 85.3%; Score 81; DB 5; Length 80;
 Best Local Similarity 89.5%; Pred. No. 3.5e-05; Indels 0; Gaps 0;
 Matches 17; Conservative 0; MisMatches 2; Indels 0; Gaps 0;

Qy	2 EKKIARMEKASSVTVNVS 20		
Db	62 EKKIKRMKEKCSVTVNVS 80		

RESULT 4

Q9U0P7	PRELIMINARY;	PRT;	80 AA.
ID			
AC			
DT	01-MAY-2000 (TREMBrel. 13; Created)		
DT	01-MAY-2000 (TREMBrel. 13; Last sequence update)		
DT	01-JUN-2003 (TREMBrel. 24; Last annotation update)		
DE	Circumsporozoite protein (Fragment).		
GN	OS		
OC	Plasmodium falciparum		
OX	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
RN	[1] NCBI_TaxID=5833;		
RP	SEQUENCE FROM N.A.		
RC	STRAN=D4230;		
RA	de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.; "Sequence variation in the non-repeat region of the Plasmodium falciparum Glutamate rich protein (GLURP) from Brazil, Senegalese, and Burmese field isolates and from laboratory strains"; Submitted (SBP-1999) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AJ269961; CAB64180.1; -.		
DR	GO; GO:0003793; F:defense/immunity protein activity; IEA.		
DR	InterPro; IPR003067; Crampsprotoite.		
DR	InterPro; IPR000884; TSP1.		
DR	Pfam; PF00090; tsp_1; 1.		
DR	PRINTS; PRO1303; CRCMSPRZOITE.		
DR	SMART; SM00209; TSP1; 1.		
DR	PROSITE; PS50032; TSP1; 1.		
FT	NON_TER 1 80		
SQ	SEQUENCE 80 AA; 9047 MW; BA7639290DB031C3B CRC64;		

Query Match 85.3%; Score 81; DB 5; Length 80;
 Best Local Similarity 89.5%; Pred. No. 3.5e-05; Indels 0; Gaps 0;
 Matches 17; Conservative 0; MisMatches 2; Indels 0; Gaps 0;

Qy	2 EKKIARMEKASSVTVNVS 20		
Db	62 EKKIKRMKEKCSVTVNVS 80		

RESULT 5

Q9TW97	PRELIMINARY;	PRT;	80 AA.
ID			
AC			
DT	01-MAY-2000 (TREMBrel. 13; Created)		
DT	01-MAY-2000 (TREMBrel. 13; Last sequence update)		
DT	01-JUN-2003 (TREMBrel. 24; Last annotation update)		
DE	Circumsporozoite protein (Fragment).		
GN	OS		
OC	Plasmodium falciparum		
OX	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
RN	[1] NCBI_TaxID=5833;		
RP	SEQUENCE FROM N.A.		
RC	STRAN=D4403; and D4259;		
RA	de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.; "Sequence variation in the non-repeat region of the Plasmodium falciparum Glutamate rich protein (GLURP) from Brazil, Senegalese, and Burmese field isolates and from laboratory strains"; Submitted (SPB-1999) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AJ269968; CAB64187.1; -.		
DR	InterPro; IPR003067; Crampsprotoite.		
DR	InterPro; IPR000884; TSP1.		
DR	Pfam; PF00090; tsp_1; 1.		
DR	PRINTS; PRO1303; CRCMSPRZOITE.		
DR	SMART; SM00209; TSP1; 1.		
DR	PROSITE; PS50032; TSP1; 1.		
FT	NON_TER 1 80		
SQ	SEQUENCE 80 AA; 9123 MW; 4614EECC68F0B1434 CRC64;		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 10, 2004, 08:58:54 ; Search time 19.7665 Seconds
Perfect score: 95 (without alignments)
Sequence: US-09-848-834A-3 DEKKTAKMEKASSVENVNS 20

Scoring table: BioSUM62 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_251,*

1: sp_bacteria;*

2: sp_fungi;*

3: sp_invertebrate;*

4: sp_human;*

5: sp_invertebrate;*

6: sp_mammal;*

7: sp_mhc;*

8: sp_organelle;*

9: sp_phage;*

10: sp_plant;*

11: sp_rabbit;*

12: sp_virus;*

13: sp_vertebrate;*

14: sp_unclassified;*

15: sp_tvirus;*

16: sp_bacteriophage;*

17: sp_archaea;*

Q9u0p8	Plasmodium	80	5
Q9top4	Plasmodium	80	5
Q9tm9	Plasmodium	80	5
Q9tw3	Plasmodium	80	5
Q9tw83	Plasmodium	80	5
Q25837	Plasmodium	115	5
Q25839	Plasmodium	115	5
Q9u34	Plasmodium	115	5
Q25836	Plasmodium	115	5
Q25835	Plasmodium	115	5
Q9tw7	Plasmodium	115	5
Q25795	Plasmodium	117	5
Q25797	Plasmodium	117	5
Q25796	Plasmodium	117	5
Q25794	Plasmodium	117	5
Q8i95	Plasmodium	117	5
Q8i97	Plasmodium	117	5
Q8i96	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i918	Plasmodium	117	5
Q8i919	Plasmodium	117	5
Q8i94	Plasmodium	117	5
Q8i93	Plasmodium	117	5
Q8i98	Plasmodium	117	5
Q8i91	Plasmodium	117	

Qy 2 HWSYGLRP 9
 Db 23 HWSYGLSP 30

RESULT 14

A21114
 gonadotropin - chum salmon
 C;Species: *Oncorhynchus keta* (chum salmon)
 C;Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
 C;Accession: A21114
 R;Sherwood, N.; Biden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
 Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
 A;Title: Characterization of a teleost gonadotropin-releasing hormone.
 A;Reference number: A21114; MUID:83195140; PMID:6341999

A;Accession: A21114

A;Status: Preliminary

A;Molecule type: protein

A;Residues: 1-10 <SHE>

Query Match 72.2%; Score 39; DB 2; Length 10;
 Best Local Similarity 75.0%; Pred No. 0.33;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 HWSYGLRP 9
 Db 2 HWSYGLP 9

RESULT 15

I51092
 gonadotropin releasing hormone - chinook salmon (fragment)
 C;Species: *Oncorhynchus tshawytscha* (chinook salmon)
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Aug-1997
 C;Accession: I51092
 R;Klungland, H.; Lorenz, J.B.; Andersen, O.; Kisen, G.O.; Alestrom, P.
 Mol. Cell. Endocrinol. 84, 167-174, 1992
 A;Title: The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA.
 A;Reference number: I51040; MUID:92267241; PMID:1587389

A;Accession: I51092

A;Status: Preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-74 <KLU>

A;Cross-references: EMBL:X79711; NID:9499322; PID:9499323

C;Genetics: A;Gene: GnRH

A;Introns: 38/3; 65/3

Query Match 72.2%; Score 39; DB 2; Length 74;
 Best Local Similarity 75.0%; Pred No. 2.9;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 HWSYGLRP 9
 Db 17 HWSYGLP 24

Search completed: March 10, 2004, 09:16:38
 Job time : 4.17166 secs

Best Local Similarity 87.5%; Pred. No. 0.088; Mismatches 0; Indels 0; Gaps 0;

RESULT 10
 RHD1S
 gonadotropin I precursor - Sharpooth catfish
 N;Alternate names: gonadotropin-releasing hormone I (GnRH-I);
 C;Contains: Gonadotropin I; gonadotropin I-associated protein form I; gonadotropin I
 C;Species: Clarias gariepinus (sharpooth catfish)
 C;Date: 30-Sep-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
 C;Accession: S45602; S45601; JJC1242; S42316; S42317;
 R;Bogerd, J.; Zandbergen, T.; Anderson, E.; Goos, H.
 Bur. J. Biochem. 222, 541-549, 1994
 A;Title: Isolation, characterization and expression of cDNAs encoding the catfish-type
 A;Reference number: S45600; MUID:94291651; PMID:8020492
 A;Molecule type: mRNA
 A;Residues: 1-80 <BOG1>
 A;Cross-references: EMBL:X78049; PIDN:CAA54971_1; PID:9459434
 A;Note: Gonadotropin I-associated protein form I
 A;Accession: S45601
 A;Molecule type: mRNA
 A;Residues: 1-46; S' 48-59; 'G' 61-80 <BOG2>
 A;Cross-references: EMBL:X78048; PIDN:CAA5470_1; PID:9459332
 A;Note: Gonadotropin I-associated protein form II, presumed to be a polymorphic form
 R;Bogerd, J.; Li, K.W.; Janssen-Dommeholt, C.; Goos, H.
 Biochem. Biophys. Res. Commun. 187, 127-134, 1992
 A;Title: Two gonadotropin-releasing hormones from African catfish (Clarias gariepinus).
 A;Reference number: JJC1242; MUID:9239313; PMID:1520292
 A;Accession: JJC1242
 A;Molecule type: protein
 A;Residues: 22-31 <BOG3>
 A;Experimental source: brain
 C;Superfamily: Gonadotropin I-associated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 C;Keywords: amidated carboxyl end; status predicted
 C;KeyWords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F;1-21/Domain: Signal sequence #status predicted <SIG>
 F;22-31/Domain: Gonadotropin I-associated protein #status predicted <MAT1>
 F;35-80/Product: gonadotropin I-associated protein #status predicted <MAT2>
 F;22/Modified site: pyrrolidine carboxylic acid (Gln) (in mature form) #status experim
 F;31/Modified site: amidated carboxyl end (Gly) (amide in mature form) from following 91
 Query Match Score 46; DB 2; Length 98;
 Best Local Similarity 87.5%; Pred. No. 0.22; Mismatches 1; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query 2 HWSYGLP 9
 QY ||||| |
 Db 24 HWSYGLSP 31

RESULT 11
 E95361
 probable muconate cycloisomerase (EC 5.5.1.1) [imported] - Sinorhizobium meliloti (strain C;Species: Sinorhizobium meliloti
 C;Accession: E95361
 R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barlow-Hubler, F.; Bowe ; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot A;Reference number: A95222; MUID:21396509; PMID:11481432
 A;Accession: E95361
 A;Status: preliminary
 A;Molecule type: DNA
 A;Cross-references: GB:AB006469; PIDN:AAK65455_1; PID:gi14523923; GSPPDB:GN00165
 A;Experimental source: strain 1021, megaplasmid pSymA
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Phphler, A.; Ampe, F.; Barlow-Hubler, R.; Pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federpel, N.A.; Fisher, R.P.; L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelauré, heault, P.; Vandembol, M.; Vorhofer, F.J.; Weider, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A;Reference number: A96039; MUID:21368234; PMID:11474104
 A;Contents: annotation
 C;Genetics: SmaI461
 A;Gene: plasmid
 C;Keywords: intramolecular lyase; isomerase

Query Match Score 43; DB 2; Length 374;
 Best Local Similarity 85.7%; Pred. No. 3.3; Mismatches 1; Indels 0; Gaps 0;

Query 2 HWSYGLR 8
 QY ||||| |
 Db 21 HWSYGIR 27

RESULT 13
 JC7393
 medaka-type gonadotropin-releasing hormone precursor - Japanese medaka
 C;Species: Oryzias latipes (Japanese medaka)
 C;Accession: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 20-Jan-2003
 R;Okubo, K.; Amano, M.; Yoshiura, Y.; Suetake, H.; Aida, K.
 Biochem. Biophys. Res. Commun. 276, 298-303, 2000
 A;Title: A novel form of Gonadotropin-releasing hormone in the medaka, Oryzias latipes.
 A;Reference number: JC7393
 A;Contents: Brain
 A;Accession: JC7393
 A;Molecule type: mRNA
 A;Residues: 1-91 <OKU>
 A;Cross-references: DDBJ:AB041333
 C;Comment: This protein plays the roles as a hypophysiotropic factor, and a physiologic
 C;Genetics:
 A;Gene: mdanrh
 C;Superfamily: gonadotropin
 C;Keywords: brain
 Query Match Score 42; DB 2; Length 91;
 Best Local Similarity 75.0%; Pred. No. 1.1; Mismatches 1; Indels 0; Gaps 0;

RiTan, L.; Rousseau, P. *Commun.* 109, 1061-1071, 1982
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982

A;Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in A;Reference number: A0108; MUID:83126573; PMID:6760865

A;Accession: A90108

A;Molecule type: protein

A;Residues: 24-33 <TAI>

A;Experimental source: Placental trophoblasts

R;Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumister, W.; Amsterda FR;Lett. 346, 203-206, 1994

A;Title: Segmental degradation of the neuropeptide gonadotropin-releasing hormone by th A;Reference number: S5718; MUID:9428359; PMID:8013634

A;Contents: annotation; degradation pathway of synthetic hormone

C;Genetics:

A;Gene: GDB:GRH; LHRH; GRH

A;Cross-references: GDB:133746; OMIM:227200; OMIM:152760

A;Map position: 8p21-8p11.2

A;Introns: 4/7/3; 79/3

C;Function:

A;Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropin A;Note: Gonadoliberin-associated protein may have prolactin release inhibiting activity C;Superfamily: Gonadoliberin

C;Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid

F;1-23-/Domain: signal sequence #status predicted <SG>

F;24-92-/Product: progonadoliberin #status predicted <PGN>

F;24-92-/Product: gonadoliberin #status predicted <GDN>

F;17-92-/Product: gonadoliberin #status experimental <MAT>

F;17-92-/Product: gonadoliberin-associated protein #status predicted <GAP>

F;24-/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment F;33-/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 96.3%; Score 52; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.017; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0; MisMatches 0; Del 0; Ins 0;

QY 2 HWSYGLRP 9
 DD 25 HWSYGLQP 32

RESULT 7
 RHAQI
 Gonadoliberin I - American alligator
 Gonadoliberin I - American alligator mississippiensis (American alligator)
 Gonadoliberin I - American alligator mississippiensis (American alligator)

C;Species: Alligator mississippiensis (American alligator)
 C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997

R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swans Regul. Pept. 33, 105-116, 1991

A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains c A;Reference number: A60066; MUID:91352338; PMID:1882082

A;Accession: A60066

A;Molecule type: protein

A;Residues: 1-10 <LOV>

C;Superfamily: Gonadoliberin

C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F;1-/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F;10-/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 88.9%; Score 48; DB 1; Length 10;
 Best Local Similarity 87.5%; Pred. No. 0.0077; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 1; MisMatches 0; Del 0; Ins 0;

QY 2 HWSYGLRP 9
 DD 2 HWSYGLQP 9

RESULT 9
 RHRNG
 Gonadoliberin precursor - rat
 Gonadoliberin names: Gonadoliberin-associated protein (GAP); gonadotropin releasing hormon N;Contains: gonadoliberin; prolactin release-inhibiting factor

C;Species: Rattus norvegicus (Norway rat)

C;Accession: A40147; B26173; A48410

R;Bond, C.T.; Hayflick, J.S.; Seenburg, P.H.; Adelman, J.P.

Mol. Endocrinol. 3, 1257-1262, 1989

A;Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex A;Reference number: A40147; MUID:89384661; PMID:2473669

A;Accession: A40147

A;Molecule type: DNA

A;Residues: 1-92 <BON>

A;Cross-references: GB:M31670; PID:g204447; PID:AAA41264_1; PID:g204448

R;Adelman, M.J.; Mason, A.J.; Hayflick, J.S.; Seenburg, P.H.

Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986

A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot A;Reference number: A94030; MUID:86094338; PMID:2867548

A;Accession: B26173

A;Molecule type: mRNA

A;Residues: 1-92 <AD>

R;Maier, C.C.; Marchetti, B.; LeBeuf, R.D.; Blalock, J.E.

Cell Mol. Neurobiol. 12, 447-454, 1992

A;Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone A;Reference number: A48410; MUID:93105480; PMID:1468115

A;Accession: A48410

A;Molecule type: mRNA

A;Residues: 1-92 <WAI>

A;Cross-references: GB:550870; PID:g262059; PID:9262059; PID:9496326; PIDN:CAA49246_1; PID:9311612

A;Experimental source: thymus

Query Match 88.9%; Score 48; DB 2; Length 92;

A;Note: sequence extracted from NCBI backbone (NCBIN:121082; NCBIP:121083)
 C;Genetics:
 C;Introns: 47/3; 79/3

A;Description: stimulates pituitary secretion of lutropin and follitropin and inhibiting activit C;Superfamily: Gonadoliberin-associated protein may have prolactin release inhibiting activit

C;Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid

F;1-3-/Domain: signal sequence #status predicted <SG>

F;24-92-/Product: progonadoliberin #status predicted <PGN>

F;24-33-/Product: gonadoliberin #status predicted <GDN>

F;37-92-/Product: prolactin release-inhibiting factor #status predicted <PIF>

F;24-/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicte F;33-/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 96.3%; Score 52; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.017; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0; MisMatches 0; Del 0; Ins 0;

QY 2 HWSYGLRP 9
 DD 25 HWSYGLQP 32

RESULT 8
 RHAQI
 Gonadoliberin I - American alligator

C;Species: Alligator mississippiensis (American alligator)
 C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997

R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swans Regul. Pept. 33, 105-116, 1991

A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains c A;Reference number: A60066; MUID:91352338; PMID:1882082

A;Accession: A60066

A;Molecule type: protein

A;Residues: 1-10 <LOV>

C;Superfamily: Gonadoliberin

C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F;1-/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F;10-/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 88.9%; Score 48; DB 1; Length 10;
 Best Local Similarity 87.5%; Pred. No. 0.0077; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 1; MisMatches 0; Del 0; Ins 0;

QY 2 HWSYGLRP 9
 DD 2 HWSYGLQP 9

RESULT 9
 RHAQI
 Gonadoliberin I Precursor - chicken

N;Alternate names: gonadotropin-releasing hormone I

C;Species: Gallus gallus (chicken)

C;Accession: 150644; S33307

R;Dunn, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.

J. Mol. Endocrinol. 11, 19-29, 1993

A;Title: Characterization of the chicken preprogonadotrophin-releasing hormone-I gene. A;Reference number: 150644; MUID:94059355; PMID:1902055

A;Accession: 150644

A;Molecule type: DNA

A;Residues: 1-92 <DU2>

A;Cross-references: EMBL:X69491; NID:9496326; PIDN:CAA49246_1; PID:9311612

C;Genetics:
 A;Introns: 47/3; 79/3

C;Superfamily: Gonadoliberin

Query Match 88.9%; Score 48; DB 2; Length 92;

A;Residues: 1-10 <BUR>
 A;Note: the natural and synthetic hormones have the same biological activity
 C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and f

C;Superfamily: gonadotropin-releasing hormone
 C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F;1;Modified site: pyrrolidine carboxylic acid (Gln) #status experimental
 F;1;Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 96.3%; Score 52; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0015; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0; A;Reference number: A47578;
 A;Accession: A47578

QY 2 HWSYGLRP 9
 DB 2 HWSYGLRP 9

QY 2 HWSYGLRP 9
 DB 7 HWSYGLRP 14

Query Match 96.3%; Score 52; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.012; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.012; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.012; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.012; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.012; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.012; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.012; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.012; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.012; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

RESULT 5

RHUG gonadotropin precursor - mouse
 N;Alternate names: Gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing

N;Contains: gonadotropin-associated protein (GAP)
 C;Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999

C;Accession: A47578
 R;Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolic
 Science 234, 1366-1371, 1986
 A;Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible f

A;Reference number: A47578;
 A;Accession: A47578
 A;Molecule type: DNA
 A;Residue: 1-90 <MAS>
 A;Cross-references: EMBL: M14872; PIDN: AAA37717.1; PID: g387175

C;Introns: 45/3; 77/3

C;Function:

A;Description: Gonadotropin stimulates pituitary secretion of lutropin and follitropin

A;Note: Gonadotropin-associated protein may have prolactin release inhibiting activity
 C;Superfamily: gonadotropin
 C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F;1;3/Domain: signal sequence #status predicted <SG>
 F;22/31/Product: gonadotropin #status predicted <GB>
 F;35/90/Product: gonadotropin-associated protein #status predicted <GAP>

F;22/Modified site: amidated carboxyl end (Gly) (amide in mature form) #status predicted
 F;31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following G)

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.016; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.016; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.016; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.016; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.016; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.016; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.016; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.016; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.016; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 DB 23 HWSYGLRP 30

A;Cross-references: GB:X01059; PID: g343556
 A;Experimental source: placenta

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using EW model

Run on: March 10, 2004, 08:58:54 ; Search time 3.11284 Seconds
3.09.015 Million cell updates/sec
(without alignments)

Title: US-09-848-834A-1
Perfect score: 54
Sequence: 1 XHNSYGLRFX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283365 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : PIR78:
1: pir1:
2: pir2:
3: pir3:
4: pir4:
* * * *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	96.3	10	RHPGG	gonadoliberin - pi
2	52	96.3	10	RHSHG	gonadoliberin - sh
3	52	96.3	67	I78541	gonadoliberin prec
4	52	96.3	89	I51423	gonadoliberin prec
5	52	96.3	90	RHMSG	gonadoliberin prec
6	52	96.3	92	RHHUG	gonadoliberin prec
7	52	96.3	92	RHRTG	F10/Modified site: amidated carboxylic acid (Gly) #status experimental
8	48	88.9	10	RHAQI	gonadoliberin I -
9	48	88.9	10	I50644	gonadotropin-I pr
10	46	85.2	98	I50739	gonadotropin-releasable muconate
11	43	79.6	374	E95361	gonadoliberin I dr
12	42	77.8	80	RHDLS	gonadotropin-I dr
13	42	77.8	91	JCT393	medaka-type Gonado
14	39	72.2	10	A21114	gonadoliberin - ch
15	39	72.2	74	I51092	gonadotropin releas
16	39	72.2	82	I51180	gonadotropin releas
17	39	72.2	82	I51355	gonadotropin-releas
18	39	72.2	82	I51365	gonadotropin-releas
19	39	72.2	90	I51331	salmon-type gonado
20	39	72.2	90	A23735	gonadotropin releas
21	39	72.2	90	I51095	gonadotropin prec
22	39	72.2	90	C72710	gonadotropin prec
23	39	72.2	388	T18856	probable fm Prote
24	39	72.2	1444	I5161	angiogenesis inhib
25	38	70.4	161	D84472	hypothetical prote
26	38	70.4	293	G72699	hypothetical prote
27	38	70.4	501	T32848	probable dehydroge
28	37	68.5	270	F88035	dehydrogenase
29	37	68.5	345	A58519	probable oxidoreduc

ALIGNMENTS

RESULT 1

RHPGG

Gonadoliberin - pig
C;Species: Sus scrofa domesticus (domestic pig)
C;Date: 13-Jul-1991 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997

C;Accession: A01411

R;Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A;Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of t

A;Reference number: A90172; MUID:72114303; PMID:4946067

A;Accession: A01411

A;Molecule type: protein

A;Residues: 1-10 <HAB>

R;Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A;Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase meth

A;Reference number: A90176; MUID:72065376; PMID:4942726

A;Contents: annotation, synthesis

A;Note: the synthetic and natural hormones have the same physicochemical and biological
R;Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A;Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A;Accession: A90175; MUID:72117544; PMID:4946275

A;Contents: annotation

A;Note: TRP-3 appears to be essential for biological activity
A;Title: This hypothalamic hormone stimulates the secretion of both luteinizing and f

C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and f

C;Superfamily: Gonadoliberin
C;Keywords: amidated carboxylic acid; hormone; hypothalamus; pyroglutamic acid
F1/F1/Modified site: Pyrrolidine carboxylic acid (Gly) #status experimental
F10/F10/Modified site: amidated carboxylic acid (Gly) #status experimental

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Prec. No. 0.0015;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSIGLRP 9
Db 2 HWSIGLRP 9

RESULT 2

RHSHG

Gonadoliberin - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C;Accession: A93780; A01411

R;Burgs, R.; Butcher, M.; Anoss, M.; Ling, N.; Monahan, M.; Fellows, R.; E
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A;Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fact

A;Reference number: A93780; MUID:72094314; PMID:4550508
A;Accession: A93780
A;Molecule type: protein

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Larson and Taylor
 STREET: 727 Twenty-Third Street, South
 CITY: Arlington
 STATE: Virginia
 COUNTRY: USA
 ZIP: 22202

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/343,883
 FILING DATE: 17-NOV-1994
 CLASSIFICATION: 424
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/946,495
 FILING DATE: 09-NOV-1992
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: FR 91.02513
 FILING DATE: 01-MAR-1991
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: FR 9115289
 FILING DATE: 10-DEC-1991
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 TOPOLogy: linear
 MOLECULE TYPE: Peptide

FEATURE:
 NAME/KEY: Peptide
 LOCATION: 10
 OTHER INFORMATION: /label= NH2
 /note= "amidated glycine"

FEATURE:
 NAME/KEY: Peptide
 LOCATION: 1
 OTHER INFORMATION: /label= PYRO
 /note= "pyroglutamic acid"
 PUBLICATON INFORMATION:
 AUTHORS: Matsuo, H.
 AUTHORS: Baba, Y.
 AUTHORS: G. Nair, R. M.
 AUTHORS: Arimura, A.
 AUTHORS: Schally, A. V.
 TITLE: Structure of the Porcine LH- and
 FSH-releasing hormone. I. The proposed amino acid
 sequence.
 JOURNAL: Biochem. Biophys. Res. Commun.
 VOLUME: 43
 ISSUE: 6
 PAGES: 1334-1339
 DATE: 1971

RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 10
 US-08-343-883-1

Query Match 96.3%; Score 52; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0025; Matches 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HWSYGLRP 9
 Db |||||
 2 HWSYGLRP 9

Search completed: March 10, 2004, 09:28:50
 Job time : 4.8716 sec

NAME/KEY: Modified-site
 LOCATION: 1
 OTHER INFORMATION: /note= "Xaa at position 1 is a
 US-08-103-022-1 5'-oxo-prolyl aminoacyl residue."

Query Match 96.3%; Score 52; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;

Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

RESULT 13
 US-07-897-680-1
 Sequence 1, Application US/07897680
 Patent No. 5445025

GENERAL INFORMATION:
 APPLICANT: Subba Rao, Gowdahallin N.
 ATTORNEY/AGENT INFORMATION:
 Fu Lu, Mou-Ying Lee, Dennis Y.

TITLE OF INVENTION: Formulations and Method for the Percutaneous Administration of Leuproreotide

NUMBER OF SEQUENCES: 2

CURRENT APPLICATION DATA:
 CORRESPONDENCE ADDRESS:
 ADDRESS: Dept. 377 Abbott Laboratories
 STREET: One Abbott Park Road
 CITY: Abbott Park
 STATE: IL
 COUNTRY: USA
 ZIP: 60064-3500

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/897,680
 FILING DATE: 19920612
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Janssen, Jitry P.
 REFERENCE/DOCKET NUMBER: 5165.US.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (708) 937-7556
 TELEFAX: (708) 938-7742

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Peptide
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 1
 OTHER INFORMATION: /note= "XAA at position 1 is a
 OTHER INFORMATION: Pyro-glutamyl residue"

LOCATION: 10
 OTHER INFORMATION: /note= "XAA at position 10 is a
 OTHER INFORMATION: glycyl-amide residue"

Query Match 96.3%; Score 52; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;
 Gaps 0;

Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

RESULT 14
 US-08-184-935-6
 Sequence 6, Application US/08184935
 Patent No. 547670
 GENERAL INFORMATION:
 APPLICANT: PRADELLES, PHILIPPE
 TITLE OF INVENTION: IMMUNOMETRIC DETERMINATION OF AN ANTIGEN
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
 ADDRESS: P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/184,935
 FILING DATE: 24-JAN-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Oblon, No. 547670man P.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 846-286-0

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELEX: 24885 OPAT UR
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 10
 OTHER INFORMATION: /note= "C-terminal amide"

US-08-184-935-6

Query Match 96.3%; Score 52; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

RESULT 15
 US-08-343-883-1
 Sequence 1, Application US/08343883
 Patent No. 5573767

GENERAL INFORMATION:
 APPLICANT: Dufour, Raymond J.
 APPLICANT: Roulet, Claude J.M.
 APPLICANT: Chouvet, Claire D.
 APPLICANT: Bonneau, Michel B.
 TITLE OF INVENTION: Method for improving the organoleptic qualities of meat from uncastrated male domestic animals, vaccines which are usable in this method, new peptide, in particular for producing these vaccines . . .
 NUMBER OF SEQUENCES: 2

Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

US-07-690-983D-32
Sequence 32, Application US/07690983D

PATENT NO. 540586

GENERAL INFORMATION:

APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
STREET: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07-690, 983D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-3399

INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-07-690-983D-32

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0025; Mismatches 0; Indels 0; Gaps 0;
Qy 2 HWSYGLRP 9
Db 2 HWSYGLRP 9

RESULT 11
Sequence 37, Application US/07690983D
PATENT NO. 540386
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
STREET: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103, 022
FILING DATE: 05-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Janssen, Jerry F.
REGISTRATION NUMBER: 29,175
REFERENCE/DOCKET NUMBER: 5389.US.O1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 938-7742
TELEFAX: (708) 938-2623
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
FEATURE:

APPLICATION NUMBER: US/07-690, 983D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-3399
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-07-690-983D-37

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0025; Mismatches 0; Indels 0; Gaps 0;
Qy 2 HWSYGLRP 9
Db 3 HWSYGLRP 10

RESULT 12
US-08-103-022-1
Sequence 1, Application US/08103022
PATENT NO. 541390
GENERAL INFORMATION:
APPLICANT: Haviv, Fortuna
APPLICANT: Fitzpatrick, Timothy D.
APPLICANT: Swenson, Rolf E.
APPLICANT: Nichols, Charles J.
APPLICANT: Mort, Nicholas A.
TITLE OF INVENTION: N-Terminal Modified Analogs of LHRH
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman, Jr., Dept. 377
STREET: Abbott Laboratories, One Abbott Park Road
CITY: No. 541390th Chicago
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103, 022
FILING DATE: 05-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Janssen, Jerry F.
REGISTRATION NUMBER: 29,175
REFERENCE/DOCKET NUMBER: 5389.US.O1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 938-7742
TELEFAX: (708) 938-2623
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
FEATURE:

FILING DATE: 25-JUN-1991
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: PCT/AU90/00373
 FILING DATE: 24 AUG-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 10
 OTHER INFORMATION: /note= "Represents glycinamide"
 US-07-690-983D-6

Query Match 96.3%; Score 52; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Missmatches 0;

Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

RESULT 9
 US-07-690-983D-8
 Sequence 8, Application US/07690983D
 Patent No. 5403586
 GENERAL INFORMATION:
 APPLICANT: RUSSELL-JONES, Gregory J.
 APPLICANT: STEWART, Andrew G.
 APPLICANT: TSONIS, Con G.
 TITLE OF INVENTION: FUSION PROTEINS
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W.
 CITY: Washington, D.C.
 COUNTRY: USA
 ZIP: 20007-5109

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/690,983D
 FILING DATE: 25-JUN-1991
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/AU90/00373
 FILING DATE: 24-AUG-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 16786/148 CHAC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399

INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 10
 OTHER INFORMATION: /note= "Represents glycinamide"

US-07-690-983D-8

Query Match 96.3%; Score 52; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Missmatches 0;

Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

RESULT 10

Query Match 96.3%; Score 52; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0;

Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

RESULT 5
 US-07-690-983D-1
 Sequence 1, Application US/07690983D

GENERAL INFORMATION:
 APPLICANT: RUSSELL-JONES, Gregory J.
 APPLICANT: STEWART, Andrew G.
 APPLICANT: TSONIS, Con G.
 TITLE OF INVENTION: FUSION PROTEINS
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W.
 CITY: Washington, D.C.
 COUNTRY: USA
 ZIP: 20007-5109

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/690, 983D
 FILING DATE: 25-JUN-1991
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: PCT/AU90/00373
 FILING DATE: 24-AUG-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 16786/148 CHAC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein

US-07-690-983D-2

Query Match 96.3%; Score 52; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

RESULT 7
 US-07-690-983D-6
 Sequence 6, Application US/07690983D

GENERAL INFORMATION:
 PATENT NO. 5403586
 APPLICANT: RUSSELL-JONES, Gregory J.
 APPLICANT: STEWART, Andrew G.
 APPLICANT: TSONIS, Con G.
 TITLE OF INVENTION: FUSION PROTEINS
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W.
 CITY: Washington, D.C.
 ZIP: 20007-5109

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/690, 983D

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.3.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/572,339
 FILING DATE: 18-May-2000
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/474,555
 FILING DATE: 07-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: SAXE, Bernhard D.
 REGISTRATION NUMBER: 28,665
 REFERENCE/DOCKET NUMBER: 18733/584/TMIN
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE: Modified-site
 NAME/KEY: Modified-site
 LOCATION: 1
 OTHER INFORMATION: /product= "OTHER"
 /note= "The Xaa at position 1 = pyroglutamic acid."
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-572-339-1

Query Match Score 96.3%; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

RESULT 3
 US-07-714-540-9
 Sequence 9, Application US/0714540
 Patent No. 5262321
 GENERAL INFORMATION:
 APPLICANT: Altintas, Ronald G.
 ATTORNEY/AGENT INFORMATION:
 APPLICANT: Toll, Lawrence
 TITLE OF INVENTION: ISOLATED ATRIAL PEPTIDE-DEGRADING ENZYME AND NOVEL COMPOUNDS USEFUL AS INHIBITORS THEREOF
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Irell & Manella
 STREET: 545 Middlefield Road, Suite 200
 STATE: Menlo Park
 COUNTRY: USA
 ZIP: 94025
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/714,540
 FILING DATE: 19-Mar-2000
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Reed, Diane E.
 REGISTRATION NUMBER: 31,292
 REFERENCE/DOCKET NUMBER: 8500-0135.00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-327-7350
 TELEFAX: 415-327-2951

Query Match Score 96.3%; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0025; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9

RESULT 4
 US-07-983-111-1
 Sequence 1, Application US/07983111
 Patent No. 5284657
 GENERAL INFORMATION:
 APPLICANT: Fu, Lu, Mou-Ying L.
 ATTORNEY/AGENT INFORMATION:
 APPLICANT: Reliland, Thomas L.
 TITLE OF INVENTION: "Compositions and Methods for the Sublingual or Buccal Administration of Therapeutic Agents"
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Edward H. Gorman, Jr., Dept. 377-AP6D
 STREET: One Abbott Park Road
 CITY: Abbott Park
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60064-3500
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/983,111
 FILING DATE: 30-NOV-1992
 CLASSIFICATION: 424
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/07/750,843
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Janssen, Jerry F.
 REGISTRATION NUMBER: 29,175
 PRIORITY/DOCKET NUMBER: 4848-US-01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (708) 938-7742
 TELEFAX: (708) 937-5556
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 1
 OTHER INFORMATION: /note= "Xaa at position 1 is 5-oxo-proline"
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 10
 OTHER INFORMATION: /note= "Xaa at position 10 is glycaminamide"
 OTHER INFORMATION:
 US-07-983-111-1

GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.

Copyright - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 3.8716 seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-1
Perfect score: 54
Sequence: 1 XHWSYGRPX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:

389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cggn2_6/ptodata/2/iaas/5A_COMBO.pep:
2: /cggn2_6/ptodata/2/iaas/5B_COMBO.pep:
3: /cggn2_6/ptodata/2/iaas/6A_COMBO.pep:
4: /cggn2_6/ptodata/2/iaas/6B_COMBO.pep:
5: /cggn2_6/ptodata/2/iaas/PCTUS_COMBO.pep:
6: /cggn2_6/ptodata/2/iaas/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	96.3	9	4 US-09-890-650-29	Sequence 29, Appl
2	52	96.3	10	1 US-09-570-339-1	Sequence 1, Appl
3	52	96.3	10	1 US-07-714-540-9	Sequence 9, Appl
4	52	96.3	10	1 US-07-983-111-1	Sequence 1, Appl
5	52	96.3	10	1 US-07-690-983D-1	Sequence 1, Appl
6	52	96.3	10	1 US-07-690-983D-2	Sequence 2, Appl
7	52	96.3	10	1 US-07-690-983D-6	Sequence 6, Appl
8	52	96.3	10	1 US-07-690-983D-7	Sequence 7, Appl
9	52	96.3	10	1 US-07-690-983D-8	Sequence 8, Appl
10	52	96.3	10	1 US-07-690-983D-32	Sequence 32, Appl
11	52	96.3	10	1 US-07-690-983D-37	Sequence 37, Appl
12	52	96.3	10	1 US-08-103-022-1	Sequence 1, Appl
13	52	96.3	10	1 US-07-897-680-1	Sequence 2, Appl
14	52	96.3	10	1 US-08-184-935-6	Sequence 6, Appl
15	52	96.3	10	1 US-08-143-883-1	Sequence 1, Appl
16	52	96.3	10	1 US-08-000-931-5	Sequence 5, Appl
17	52	96.3	10	1 US-08-428-488-12	Sequence 22, Appl
18	52	96.3	10	1 US-08-341-219-11	Sequence 11, Appl
19	52	96.3	10	1 US-08-453-588-2	Sequence 2, Appl
20	52	96.3	10	1 US-08-453-588-4	Sequence 4, Appl
21	52	96.3	10	1 US-08-453-588-6	Sequence 6, Appl
22	52	96.3	10	1 US-08-453-588-8	Sequence 8, Appl
23	52	96.3	10	1 US-08-453-888-10	Sequence 10, Appl
24	52	96.3	10	1 US-08-453-888-12	Sequence 12, Appl
25	52	96.3	10	1 US-08-453-588-14	Sequence 14, Appl
26	52	96.3	10	1 US-08-453-588-16	Sequence 16, Appl
27	52	96.3	10	1 US-08-453-588-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1	US-09-890-650-29	Query Match %	96.3%	Score 52;	DB 4;	Length 9;
		; Sequence 29, Application US/09890650				
		; Patent No. 668547				
		; GENERAL INFORMATION:				
		; APPLICANT: JACKSON, DAVID CHARLES				
		; APPLICANT: SOUAVI, GHOSH				
		; TITLE OF INVENTION: T HELPER CELL EPITOPE				
		; FILE REFERENCE: 47-152				
		; CURRENT APPLICATION NUMBER: US/09/890,650				
		; CURRENT FILING DATE: 2001-11-05				
		; NUMBER OF SEQ ID NOS: 30				
		; SOFTWARE: PatentIn Ver. 2.1				
		; SEQ ID NO: 29				
		; LENGTH: 9				
		; TYPE: PRT				
		; ORGANISM: Canis sp.				
		; US-09-890-650-29				

RESULT 2

Query Match % 96.3% ; Score 52; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 3e+05; Missmatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 8; Number of SEQ ID NOS: 30

Qy 2 HWSYGRPX 9
Db 1 HWSYGLRP 8

Query Match % 96.3% ; Score 52; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 3e+05; Missmatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 8; Number of SEQ ID NOS: 30

Qy 2 HWSYGLRP 9
Db 1 HWSYGLRP 8

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

APPLICANT: MCBRIDE, William J.
APPLICANT: KARACI, Habibe
APPLICANT: GRIFFITHS, Gary L.

GENERAL INFORMATION:

ADDRESS: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20007-5109

TITLE OF INVENTION: RADIONETAL-BINDING ANALOGUES OF LEUTINIZING HORMONE RELEASING HORMONE

```

SEQ ID NO 3
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Gonadotrophin
OTHER INFORMATION: Release Hormone Analogue
NAME/KEY: MOD_RES
LOCATION: (1)
OTHER INFORMATION: Xaa at position 1 is PyroGlu; PYRROLIDONE
OTHER INFORMATION: CARBOXYLIC ACID
NAME/KEY: MOD_RES
LOCATION: (10)
OTHER INFORMATION: Xaa at position 10 is ethylamide;;
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 5631226
PATENT FILING DATE: 1993-07-07
PUBLICATION DATE: 1997-05-20
PUBLICATION INFORMATION:
AUTHORS: Rousselie, Christophe
JOURNAL: Mol. Pharmacol.
VOLUME: 57
PAGES: 679-686
DATE: 2000
US-09-B10-601-3

Query Match 96.3%; Score 52; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.022; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; MisMatche 0; Indels 0; Gaps 0;

RESULT 7
US-09-964-201A-28
Sequence 28, Application US/09964201A
; Publication No. US20030091575A1
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, April L
; APPLICANT: Loinas, Gerald L
; APPLICANT: Roberts, Steven F
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; CURRENT APPLICATION NUMBER: US/0991575A1
; CURRENT FILING DATE: 2002-05-21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Porcine
; NUMBER OF SEQ ID NOS: 35
; US-09-964-201A-28

Query Match 96.3%; Score 52; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.022; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; MisMatche 0; Indels 0; Gaps 0;

RESULT 8
US-09-964-201A-32
Sequence 32, Application US/09964201A
; Publication No. US20030091575A1
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, April L
; APPLICANT: Loinas, Gerald L
; APPLICANT: Roberts, Steven F
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; CURRENT APPLICATION NUMBER: US/0991575A1
; CURRENT FILING DATE: 2002-05-21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Porcine
; NUMBER OF SEQ ID NOS: 35
; US-09-964-201A-32

Query Match 96.3%; Score 52; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.022; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; MisMatche 0; Indels 0; Gaps 0;

RESULT 9
US-09-964-201A-33
Sequence 33, Application US/09964201A
; Publication No. US/0991575A1
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, April L
; APPLICANT: Loinas, Gerald L
; APPLICANT: Roberts, Steven F
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; CURRENT APPLICATION NUMBER: US/0991575A1
; CURRENT FILING DATE: 2002-05-21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Porcine
; NUMBER OF SEQ ID NOS: 35
; US-09-964-201A-33

Query Match 96.3%; Score 52; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.022; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; MisMatche 0; Indels 0; Gaps 0;


```

US-09-305-924-1
 Sequence 1, Application US/09305924A
 GENERAL INFORMATION:
 APPLICANT: Jack G. Manns
 APPLICANT: Stephen D. Acres
 APPLICANT: Richard Harland
 TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
 FILE REFERENCE: 9001-0049
 CURRENT APPLICATION NUMBER: US/09/305,924A
 CURRENT FILING DATE: 1999-05-05
 EARLIER APPLICATION NUMBER: US 60/084,217
 EARLIER FILING DATE: 1998-05-05
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 10
 TYPE: PRT
 ORGANISM: GnrH
 FEATURE: MOD_RES
 NAME/KEY: MOD_RES
 LOCATION: (1)
 OTHER INFORMATION: Xaa is pyroglutamic acid

US-09-305-924-1

Query Match 96.3%; Score 52; DB 10; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.022;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2 HWSYGLRP 9
Db	2 HWSYGLRP 9

RESULT 10
 US-09-305-924-9
 Sequence 9, Application US/09305924A
 Publication No. US09305924A
 GENERAL INFORMATION:
 APPLICANT: Jack G. Manns
 APPLICANT: Stephen D. Acres
 APPLICANT: Richard Harland
 TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
 FILE REFERENCE: 9001-0048
 CURRENT APPLICATION NUMBER: US/09/305,924A
 CURRENT FILING DATE: 1999-05-05
 EARLIER APPLICATION NUMBER: US 60/084,217
 EARLIER FILING DATE: 1998-05-05
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 9
 LENGTH: 10
 TYPE: PRT
 ORGANISM: GnrH

US-09-305-924-9

Query Match 96.3%; Score 52; DB 10; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.022;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2 HWSYGLRP 9
Db	2 HWSYGLRP 9

RESULT 11
 US-10-109-331-2
 Sequence 2, Application US/10109331
 Publication No. US20020165159A1
 GENERAL INFORMATION:
 APPLICANT: McCann, Samuel M.
 APPLICANT: Yu, Wen H.
 TITLE OF INVENTION: FSH-Releasing Peptides

NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: John H. Runnels
 STREET: P. O. Box 2471
 CITY: Baton Rouge
 STATE: LA
 COUNTRY: USA
 ZIP: 70821-2471
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25;
 Wordperfect 5.1; No. US20020165159A1 epad Version 4.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/109,331
 FILING DATE: 28-Mar-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/297,989
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Runnels, John H.
 REGISTRATION NUMBER: 33451
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (225) 387-3221
 TELEFAX: (225) 346-8049
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-10-109-331-2

Query Match 96.3%; Score 52; DB 13.; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.022;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2 HWSYGLRP 9
Db	2 HWSYGLRP 9

RESULT 12
 US-10-184-126-1
 Sequence 1, Application US/10184126
 Publication No. US20020183257A1
 GENERAL INFORMATION:
 APPLICANT: EL TAYAR, Nabil
 APPLICANT: ZHAO, Xuan
 APPLICANT: BENTLEY, Michael D.
 TITLE OF INVENTION: PEG-LHRH ANALOG CONJUGATES
 FILE REFERENCE: EL-TAYAR-2A
 CURRENT APPLICATION NUMBER: US/10/184,126
 CURRENT FILING DATE: 2002-06-28
 PRIOR APPLICATION NUMBER: US/09/698,134
 PRIOR FILING DATE: 2000-10-30
 PRIOR APPLICATION NUMBER: 60/083,340
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: PCT/US99/09160
 PRIOR FILING DATE: 1999-04-28
 NUMBER OF SEQ ID NOS: 1
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 1
 LENGTH: 10
 TYPE: PRT
 ORGANISM: Artificial Sequence